

Enhanced Metastable Dynamic States of TAM Kinases for Inhibitor Binding

A Thesis submitted towards the partial fulfilment the Degree of

DOCTOR OF PHILOSOPHY

Submitted by

GATTA KRS NARESH

(Reg.No: 17CHPH10)

Computational Biochemistry

Under the guidance of

Prof. Lalitha Guruprasad



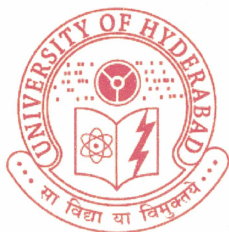
**School of Chemistry
University of Hyderabad
Hyderabad - 500046
INDIA
Jul-2023**

DEDICATION

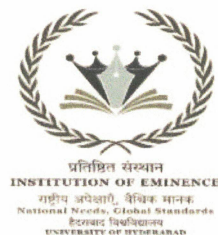
I would like to dedicate my doctoral thesis to my inspirational researcher and who's life story motivate me towards into modern advanced life sciences research field is "**Dr. Gregor Mendel**".

The man who behind the innovation of AC induction motor and modern AC electric power industry and numerous innovations which are useful to modern digital technology, I would like dedicate my doctoral thesis to the great "**Nikola Tesla**".

In my future research also I would be dedicate my research work to these two legendary scientists and futurists to change Science and technology fields into the modern technology in genomics and digital tech from 19 th century.



University of Hyderabad
School of Chemistry
Hyderabad – 500046



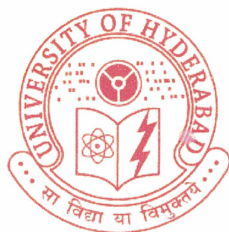
STATEMENT

I hereby declared that the matter embodied in this thesis is the result of investigation carried out by me in the School of Chemistry, University of Hyderabad, Hyderabad, under the supervision of **Prof. Lalitha Guruprasad**.

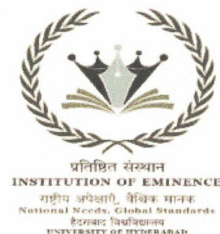
In keeping with general practice of reporting scientific observations due acknowledgements have been made whenever the work described is based on the finding of other investigators.

Hyderabad
June 2023

GATTA KRS NARESH



University of Hyderabad
School of Chemistry
Hyderabad - 500046



DECLARATION

I, **GATTA KRS NARESH** hereby declare that the thesis entitled “**Enhanced Metastable Dynamic States of TAM Kinases for Inhibitor Binding**” submitted by me under the supervision of **Prof. Lalitha Guruprasad** is a bonafide research work which is free from plagiarism. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma. I hereby agree that my thesis can be deposited in Shodhganga/INFLIBNET.

G. K. R. S. Naresh

GATTA KRS NARESH

Signature

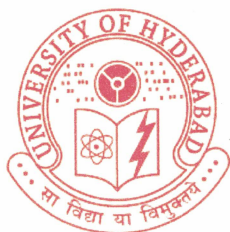
Reg. No: 17CHPH10

Lalitha 28/7/23

Prof. Lalitha Guruprasad

(Thesis Supervisor)

Lalitha Guruprasad
Professor
School of Chemistry
University of Hyderabad
Hyderabad-500 046.



University of Hyderabad
School of Chemistry
Hyderabad - 500046



CERTIFICATE

This is to certify that the thesis entitled “**Enhanced Metastable Dynamic States of TAM Kinases for Inhibitor Binding**” submitted by **Mr. GATTA KRS NARESH** holding registration number **17CHPH10** in partial fulfilment of the requirements for award of Doctor of Philosophy in the School of Chemistry is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

Parts of the thesis have been:

A. Published in the following publications

1. Naresh, G.; Guruprasad, L., Enhanced metastable state models of TAM kinase binding to cabozantinib explains the dynamic nature of receptor tyrosine kinases. *J Biomol Struct Dyn* **2021**, 39 (4), 1213-1235.
2. Naresh, G.; Guruprasad, L., Dynamic conformational states of apo, ATP & cabozantinib bound TAM kinases to differentiate active-inactive kinetic models. *J Biomol Struct Dyn* **2023**, 1-21.
3. Naresh, G.; Guruprasad, L., Mutations in the receptor-binding domain of human SARS CoV-2 spike protein increases its affinity to bind human ACE-2 receptor. *J Biomol Struct Dyn* **2023**, 41 (6), 2368-2381.
4. Guruprasad, L.; Andola, P.; Banerjee, A.; Laxman, D.; Naresh, G. K. R. S., Chapter 8 - Structure-based methods in drug design. In *Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development*, Roy, K., Ed. Academic Press: **2023**; pp 205-237.
5. In silico design, modelling and molecular mechanisms of Axl receptor tyrosine kinase inhibitors (**manuscript under preparation**)

B. List of conferences presented my Oral and Poster presentations

Posters: (ChemFest-2018); (ChemFest-2019); (ICDD-2020, 2022); (ChemFest-2022)

Orals: (MedChem 2021); (ChemFest-2022).

Further the student has passed the following courses towards fulfilment of course work requirement for Ph.D.

Course code	Name of Course	Credits	Pass/Fail
CY801	Research Proposal	4	Pass
CY805	Instrumental Methods-A	4	Pass
CY806	Instrumental Methods-B	4	Pass

Dean

School of Chemistry

Dean
SCHOOL OF CHEMISTRY
University of Hyderabad
Hyderabad-500 046.

Prof. Lalitha Guruprasad
(Thesis Supervisor)

Lalitha Guruprasad
Professor
School of Chemistry
University of Hyderabad
Hyderabad-500 046.

ACKNOWLEDGEMENTS

It is an immense gratitude to me, I would like to thank my supportive supervisor and mentor, Senior Professor “**Dr. K Lalitha Guruprasad**” for her invaluable patience and feedback, who generously provided knowledge and expertise and motivated me towards my doctoral research work. I would like to acknowledge my doctoral committee members Professor **M J Swamy** and Professor **R Nagarajan** their constant support and valuable suggestions during my doctoral thesis. It is the time to share the credit of my thesis work to my mentor for her co-operation and constant guidance and the support of my lab mates. I would like to share my special gratitude to **Dr. Guruprasad** sir for his constant guidance and valuable suggestions during research work. I would also like to thank School of Chemistry PhD students and non-teaching staff for their help wherever required. I will try to stand up to their expectation in my future research career. They were like supporting pillars throughout my doctoral program. I can’t find words to express my gratitude to my lab-mates (**Dr. Indra Neela, Dr. Shalini, Dr. Magged, Dr. Laxman, Mr. Jishu, Ms. Priyanka, Ms. Adrija, Mr. Ganesh**) for their unconditional support and excellent motivation towards my research work. I would like to give my special acknowledgement to all my doctoral 2017 batch-mates and to those who me to pursue my research career in University of Hyderabad.

I am indebted to my parents and family members for their continuous support towards my research career and understanding when undertaking research as an ambition of my life. They have always encouraged and motivated me to pursue my research goals with their unconditional love and affection. Finally, I must express my very profound gratitude to my parents **Mr. G Nookaraju & Mrs. G Kanaka Maha Lakshmi** and to my elder Sister **Mrs. G Padmavathi** and elder brother **Mr. G Ganapathi** and their family members also for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of research in scientific world. Their belief in me has kept me in high spirits and motivation during this entire process. I would also like to thank my siblings, my cousins (**Swaroop, Anjali, Pavani, Bhavani, Charan, Sai and Ajay**), my uncles (**P. Apparao, V. Bujji, V. Sathyanarayana**) and aunties (**three Laxmies**) and niece (**Poojitha, Riya**), nephew (**Yeswanth, Nanda**), son (**Shanmukh**), daughter (**Parinika**) & all my family members for their immense inspirational words and emotional support in pursuing my research interest in fundamental sciences.

I would like special acknowledge to my first mentor, who inspirational words stepping towards biotechnology is **Mr. Sudhakar** sir (B.Sc. Faculty) and **Dr. Rajesh Anna** who’s nature is fair and treat me like a younger brother and support a lot in my education.

ABBREVIATIONS

ACE-2	Angiotensin Converting Enzyme-2
ACPYPE	AnteChamber Python Parser interface
ADME	Absorption, Distribution, Metabolism, Excretion
AMBER	Assisted model building with energy refinement
ANM	Anisotropic normal modes
C spine	Catalytic spine
CHARMM	Chemistry Harvard Macromolecular Mechanics
DFG motif	Asp-Phe-Gly
EGF	Epidermal growth factor
GROMACS	GRoningen Machine for Chemical Simulations
HMM	Hidden Markov state models
JAK	Janus kinase
MD	Molecular dynamics
MEK/MAP2K	Mitogen-activated protein kinase
MSM	Markov's state model
NumPy	Numerical Python
PCA	Principal Component Analysis
PI3K	PhosphoInositide 3-Kinase
PyEMMA	Python Emma's Markov Model Algorithms
R spine	Regulatory spine
RBD	Receptor-Binding Domain
RIN	Residue interaction network
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RTKs	Receptor Tyrosine Kinases
SARS CoV-2	Severe Acute Respiratory Syndrome CoronaVirus-2
SciPy	Scientific Python
TAM	Tyro3 Mer Axl
VMD	Visual Molecular Dynamics

Amino acids

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

Contents

	Page No.
Title	-
Dedication	0
Statement	i
Declaration	ii
Certification	iii
Acknowledgements	iv
Abbreviations	v
1.0. Chapter -1 title	1
1.1. Introduction	
1.1.1. The origin of kinases	2
1.1.2. Classification of protein kinases	3
1.1.3. The receptor and non-receptor tyrosine kinases	
1.1.3.1. Receptor tyrosine kinase family	3
1.1.3.2. Non-receptor tyrosine kinase family	6
1.1.4. Mechanistic pathways of tyrosine kinases activation	6
1.2. Severe acute respiratory syndrome coronavirus 2	8
1.3. Material and methods	
1.3.1. Protein sequence databases	11
1.3.2. Multiple sequence alignment of protein sequences (MSA)	11
1.3.3. Protein homology modelling and validation	12
1.3.4. Protein data bank	13
1.3.4.1. Small molecule databases	14
1.3.4.2. Pharmacophore generation and virtual screening of chemical libraries	15
1.3.5. Computer aided drug design	16
1.3.5.1. FDA approved Kinase inhibitors and their receptor targets	18
1.3.6. Molecular docking	19
1.3.7. Absorption, distribution, metabolism,	20

excretion and toxicology studies	
1.3.8. Molecular dynamics simulations	21
1.3.8.1. Force fields	22
1.3.8.2. H-atoms addition to apo form with H++ server	24
1.3.8.3. Protein preparation for GROMACS and AMBER MD simulation	25
1.3.8.4. Minimization and NPT, NVT, Production runs of longer range MD	25
1.3.9. Basic trajectory data analysis	26
1.3.9.1. MDTraj and cpptraj, pytraj analysis	26
1.3.9.2. RMSD	27
1.3.9.3. RMSF	27
1.3.9.4. Binding free energy estimations	28
1.3.9.5. Solvent interaction energies	30
1.3.10. Python library based advanced trajectory data analysis	
1.3.10.1. Prody analysis	31
1.3.10.2. Principal component analysis	31
1.3.10.3. PyEMMA analysis	32
1.3.10.4. MSM builder	33
1.3.10.5. Markov state models	34
1.3.11. Data visualization modules	
1.3.11.1. Graphical and plotting modules	34
1.3.11.2. Reduction of dimensionality with neural networks	35
1.3.11.3. The MD traj C α data into dimensionality reduction	35
2.0. Chapter -2 title & Abstract	38
2.1. Introduction	39
2.2. Materials and methods	
2.2.1. TAM kinase structures and Homology modeling	39
2.2.2. Molecular docking	40
2.2.3. Molecular dynamics simulations	40
2.2.4. Post-MD data analysis	41

2.3. Results and Discussion	
2.3.1. Sequence analysis	43
2.3.2. Molecular docking of cabozantinib	44
2.3.3. MD simulations of apo and TAM kinase- cabozantinib complexes	47
2.3.4. Post-MD Analysis	53
2.4. Conclusions	66
3.0 Chapter -3 title	69
3.1. Introduction	70
3.2. Materials and methods	
3.2.1. Structures of apo, active and inactive TAM RTK kinase domains	71
3.2.2. Molecular docking of ATP and cabozantinib	71
3.2.3. Molecular dynamics simulations	71
3.2.4. Data analysis	72
3.3. Results and Discussion	72
3.3.1. Active – inactive kinetic state models of TAM RTKs	73
3.3.2. TAM kinase-cofactor complex activation pathway	81
3.3.3. Confirmation of the existence of active states in ATP and active/inactive states in cabozantinib bound TAM kinases	82
3.3.4. Post-MD data analysis of TAM RTK kinase domain	86
3.3.5. Kinetic transition state analysis	90
3.3.6. Mechanistic strategy of TAM RTKs activation when complexed with cabozantinib	94
3.4. Conclusions	99
4.0 Chapter -4 title & Abstract	101
4.1. Introduction	102
4.2. Materials and methods	
4.2.1. Pharmacophore featured virtual screening of chemical libraries against Axl kinase domain	102
4.2.2. Molecular docking of macromolecule (7YS) and screened-in molecules	103
4.2.3. ADME data analysis of seven hit molecules	103
4.2.4. Molecular dynamics simulations	103

4.2.5. MD simulations data analysis	104
4.2.6. Principal component analysis	104
4.3. Results and Discussion	
4.3.1. Pharmacophore model generation for virtual screening	105
4.3.2. Molecular docking of screened molecules	105
4.3.3. ADME data analysis and docking data correlations	106
4.3.4. Molecular dynamics simulations	108
4.3.4.1. Axl bound ATP and the inactive/active Axl kinases complexed with 7YS	109
4.3.5. Binding free energy calculations	
4.3.5.1. MM-GBSA (Binding affinity energy)	113
4.3.5.2. MM-PBSA (Binding Gibbs free energy)	115
4.3.6. Principal component analysis	116
4.3.7. Overall mechanism of kinase inhibition with screened small molecules	118
4.3.8. Non-bonding interactions of Axl RTKs with hinge region and regulatory site bound hit molecules	118
4.4. Conclusions	119
5.0. Chapter -5 title & Abstract	121
5.1. Introduction	122
5.2. Methods	
5.2.1. Generation of mutants in human SARS CoV-2 Spike protein RBD	124
5.2.2. Molecular dynamics simulations	125
5.2.3. Post-MD data analysis	125
5.3. Results and Discussion	
5.3.1. Molecular dynamics simulations	126
5.3.2. Ionic, hydrophobic, hydrogen bond interactions in wild-type and mutant complexes	131
5.3.3. Cluster analysis	134
5.3.4. Solvated interaction energies	134
5.4. Conclusions	138

List of references	139
List of Publications	167
List of conferences and posters, oral presentations	168
Plagiarism report and similarity index	174

Chapter -1

Introduction and Methods

1.1. Introduction

1.1.1. The origin of kinases

Kinases are a class of enzymes that belong to the transferase group and are required in order to add a phosphate group to selective substrates. Kinases play essential roles to regulate metabolic activities in the prokaryotic and eukaryotic cells (Das Evcimen et. al., 2007). Diverse classes of kinases originated from different cellular types to fulfil emerging metabolic needs in various cell types. Most classes of organisms have several families of kinases for genomic and cellular metabolic activities for cell survival from the past 3 billion years approximately (Miller et. al., 2012). The kinase phosphorylation can be on various types of substrates like histidine, serine, tyrosine, threonine in proteins, and different types of phospholipids, carbohydrates and nucleic acids. Based upon substrate specificity, the kinases are named respectively and grouped into different categories. For example, kinases that phosphorylate serine/threonine residues in proteins are called as serine/threonine kinases, kinases that phosphorylate tyrosine in proteins are called as tyrosine kinases, and histidine kinases phosphorylate histidine, lipid kinases phosphorylate lipids such as phosphatidylinositol and etc (Aehnlich et al., 2021; Graham et al., 2014). The eukaryotic organisms need to respond to their external environmental factors with the role of protein kinases to help in various cellular coordination mechanisms. These activities are initiated by the cellular transduction signaling process through the cells to trigger activation of secondary messengers like cyclic adenosine monophosphates. Kinases have evolved in a single cell prokaryotic organisms earlier than eukaryotic unicellular and multicellular organisms around 1 billion years ago (Miller et al., 2012). Protein tyrosine kinases were identified in genomic analysis of choanoflagellates, metazoan species and various other Protista species. Diverse morphological domains and large abundance of tyrosine kinases are present in many Animalia also. For example, the *Monosiga brevicollis* species have more than 128 members of tyrosine kinases and the genomic analysis studied describes these kinases to have been existed in pre-protistaera species like Opisthokonts. The sister groups of metazoans and amoeba such as *Ministeriavibrans* comprise 103 diverse kinases with their genomic data preserved in Filasterea family (Lai, & Safaei, Pelech, 2016). The multicellular level communication has evolved after this family in order to increase the size and network from metazoans and choanoflagellates subgroups. Approximately 600 million years ago, there was a rapid increase in the number of tyrosine kinase functions and therefore the encoded genes in multicellular and unicellular species have remarkably increased, new tyrosine kinases

have evolved in ecological higher order organisms. Current Animalia cells protein components have increased to suit the cellular requirements; three major components in tyrosine kinases are required to catalyze the involvement of phosphate groups; the tyrosine kinases to add the phosphate group, protein tyrosine phosphatases to remove the phosphate group and phosphorylated tyrosine binding domains such as Src homology 2 (SH2) domain.

1.1.2. Classification of protein kinases

The superfamily of protein kinases are classified based upon the modification in protein and the conformational changes due to substrate molecules binding in a reversible manner or new secondary messengers generated. Therefore, the protein kinases are classified into two types (Hunter et. al., 1985). First one is the protein kinases that undergo post-translational modifications in cellular signal transduction process due to phosphorylation of substrate molecules with the help of ATP as a co-factor. Phosphorylation is a post-translational modification, akin to methylation, glycosylation, lipidation, isoprenylation, partial proteolysis (Patterson et. al., 2014). In the second classification, the kinase activity initiated new messenger like protein molecules. According to studies (Patterson et. al., 2014; Glassman et. al., 2022), there are 568 human originated kinases in the entire proteome. It accounts for 2% of entire proteome in humans. Protein phosphorylation is performed by kinases with the help of ATP co-factor to selective hydroxyl group containing residues (serine, threonine and tyrosine) in the target protein. In general, most of the protein kinases catalyze phosphorylation of threonine or serine in substrate molecules and small group of kinases catalyze phosphorylation on tyrosine.

1.1.3. The receptor and non-receptor tyrosine kinases

1.1.3.1. Receptor tyrosine kinase family

The classification of tyrosine kinases is based upon their location in the cell. They are defined as receptor tyrosine kinases (RTKs) and non-receptor tyrosine kinases (nRTKs). The RTKs interact with many other downstream kinases to mediate physiological activities of the cell. Mainly the cellular communication is done with RTK signal transduction pathways. In biological systems, this process is quite complicated and yet synchronized during cellular signaling process with various kinase domains and regulatory proteins coordinating to decide the outcome of the cell in fundamental units of life (Glassman et. al., 2022). RTKs are the front line receptors for signal transduction processes in the entire protein kinase network. The RTKs are essential single-pass membrane spanning proteins in cellular signal transduction pathways (Lemke, 2013). All RTKs

share similar protein architecture, the amino terminal glycosylated extracellular region is stabilized by disulfide bonds followed by a membrane spanning segment, and an intracellular region that comprises a kinase domain (Graham, DeRyckere, Davies, & Earp, 2014). Many of these RTKs are generally activated by receptor-specific ligands, for example, binding of epidermal growth factor (EGF) to the extracellular domain of EGF receptor causes receptor dimerization followed by kinase activation and autophosphorylation. The kinase domain motions indirectly trigger the structure and ligand based intracellular communications in metabolic activities. RTKs initiate the cellular communication with the help of surface bound cytokines. These are responsible for regulation, differentiation of leukocytes and its memory cells of immunological reactions, RTKs are involved in regulating conventional signaling process in cellular genome expression and differentiation and proliferation and apoptosis. These events lead to intracellular ligand recognition and downstream signaling events (Sasaki et al., 2006; Tsou et al., 2014). Therefore the activation of RTKs play a vital role in controlling protein expression, regulation of normal physiological events in cell survival, proliferation, growth and death. Abnormal overexpression and mutations in RTKs initiate uncharacteristic cascade of signaling pathways which have numerous effects on the cellular role of proteins and their activities, such as disruption of normal cellular functions, malignancy in cells, and support the development and progression of different human cancers (Bosurgi et al., 2013; Rothlin, Carrera-Silva, Bosurgi, & Ghosh, 2015; Chien et al., 2016). The intracellular kinase domain in RTKs is a viable drug target for cancer treatment and several inhibitors have been designed and validated as cancer drugs (Huey, Minson, Earp, DeRyckere, & Graham, 2016; Kimani et al., 2016; Pinato, Chowdhury, & Stebbing, 2016; Schmitz et al., 2016). The upregulated RTKs have vital role in oncogenic signaling process that act as regular cells to inhibit the immunological reactions. It is an open challenge to study how RTKs are responsible for cancer and therefore to design their inhibitors in order to combat cancer (Davra, Kimani, Calianese, & Birge, 2016; Deng, Chen, & Han, 2016).

One subfamily of RTKs consists of three proteins, Tyro3, Axl and Mer collectively called as “TAM RTKs”. TAM RTKs have an extracellular N-terminal region comprising two immunoglobulin (Ig)-like domains and two fibronectin type III binding domains, a central single trans-membrane spanning helix and an intracellular region comprising the kinase domain (Wang et al., 2016). The immune response is triggered by interferon (IFN), interleukins (ILs) during inflammation and malfunction of cell metabolic activities. TAMs undergo extensive post-translational modifications such as glycosylation, phosphorylation and ubiquitination, their

molecular weights range from 100 to 140 kDa for Axl and Tyro3, and 165 to 205 kDa for Mer RTKs (Zhou et al., 2016). The important extracellular factors such as Growth arresting specific protein (Gas-6), Protein S (Pros1) and EGF, activate TAM kinases (Akalu, Rothlin, & Ghosh, 2016, Patterson et. al., 2014; Glassman et. al., 2022). These ligands bind to the extracellular domains of RTKs which leads to their dimerization followed by specific residue autophosphorylation in the cytoplasmic polypeptide (Ekyalongo, & Yee, 2017). TAMs play crucial roles in a variety of normal biological functions such as spermatogenesis, bone physiology, controlling platelet aggregation, endothelial and vascular smooth-muscle homeostasis (Gay, Balaji, & Byers, 2017; Vouri, & Hafizi, 2017). TAMs regulate downstream signalling events are mediated by JAK, p38, MEK, phospholipase C (PLC) and PI3K that essentially play role in cell growth, apoptosis and survival (Graham, DeRyckere, Davies, & Earp, 2014). It has been proposed that TAM kinases are emerging as a class of innate immune checkpoints that participate in key steps of anti-tumoral immunity (Akalu, Rothlin, & Ghosh, 2016). The TAM kinases are key transferases that exhibit distress transduction signaling process to further activate the T-helper and T-killer immune cells through caspase enzyme from apoptosis. These enzymes play an essential role in lysozyme activity of different malignant cell lines in our body. TAMs are associated with most cancers as overexpression of signaling pathways induced metastatic stage including angiogenesis formation to uncontrolled growth of malignant cells in human (Aehnlich et. al., 2021). TAM RTKs are also reported to play crucial roles in disease conditions such as acute myeloid leukemia, breast, colorectal, lung, ovarian cancers and glioblastoma (Rankin, & Giaccia, 2016; Knubel et al., 2014). Since TAMs are overexpressed in many oncogenic cells, some previously reported kinase inhibitors are experimentally validated as TAM RTK inhibitors (Wu et al., 2017; Wu et al., 2018).

The three TAM kinase members share a high degree of sequence and structural homology in their kinase domains. Cabozantinib is a small molecule inhibitor that is targeted towards multiple kinases such as Axl, c-Met, VEGFR2, RET, KIT and FLT3 (Grüllich, 2014). Cabozantinib was approved by U.S. Food and Drug Administration (FDA) for advanced renal cell carcinoma, hepatocellular carcinoma and medullary thyroid cancer. In September 2021, FDA has approved cabozantinib also for differentiated thyroid cancer that has progressed following prior VEGFR-targeted therapy. Cabozantinib is reported to bind TAM kinases with high affinity at nanomolar concentrations (Gajiwala et al., 2017, Sultan et al., 2017, Skora et al., 2013, Turner and Blythe,

2019, Herum, et al., 2017, Qin et al., 2019, Lacy et al., 2018, Pantano et al., 2016, Myers et al., 2019, Robinson, 2013).

1.1.3.2. Non-receptor tyrosine kinase family

The nRTKs are intracellular located cytosolic enzymes. These kinases also catalyze the transfer of inorganic phosphate group from biological high energy molecule (ATP) to selective target protein tyrosine residues. Intracellular protein-protein interactions activate the nRTKs. Src kinases are one of the members of nRTKs that are well studied. The Src kinases have a very different architecture compared to RTKs, but the regulation of these nRTKs have similar roles in cellular homeostasis and metabolic processes. The nRTKs mediate the cytosolic regulation activities such morphogenesis, cell motility and cell division. Mutant nRTKs are overexpressed in protein signaling processes and their defective functions in many cell types are linked to diseases such as metastatic cancers, tumor necrosis. The Src kinase has distinct structural domains, specified as N-terminal domain, Src homology 3 (SH3), SH2 domains, C-terminal domain and tyrosine kinase domain. Each of these domains perform a specific function essential to carryout normal signal transduction process. The Src kinase is involved in various cellular activities without any external ligand activation, as it is a cytosolic protein to initiate the cellular functions with the help of phosphorylation at specific residues on itself (autophosphorylation) or neighboring (substrate) proteins. The N-terminal domain has one or more acylation sites and anchored the entire Src kinase protein to the cell membrane. The SH2 domain is a key element to bind the phosphorylated site of tyrosine (for tyrosine kinases) or serine/threonine (serine/threonine kinases) residues of the target protein. Whereas the SH3 domain binds to proline rich peptide. The selective tyrosine residues are phosphorylated on target protein with the help of ATP cofactor and referred to as a “tyrosine kinase domain”.

1.1.4. Mechanistic pathways of tyrosine kinase activation

Protein kinase activation is based upon the ligand binding and various growth factors induced kinase activity. The kinases can exist as active and inactive state forms. The ATP cofactor binds to the catalytic domain of RTKs whereas the phosphorylation takes places on hydroxyl group residues of target protein. The catalytic site is defined as the nucleoside ribose sugar triphosphate of ATP binding to a cleft between the N-terminal and C-terminal domains connected with a hinge residues region (Kornev, A. P., Haste, N. M., Taylor, S. S., and Eyck, L. F. 2006). In a kinase domain. specified regions such as the phosphate binding loop; P-loop (Gly rich loop) located in

the N-terminal region, the helical region located opposite to P-loop represented as α C-helix region, the juxta membrane unfolding and refolding takes place in C-terminal domain in close association to DFG motif represented as “activation loop”. The three residues interactions of selective residues from P-loop, α C-helix, DFG motif are key salt bridges to represent kinase active and inactive states. The salt bridge existed between (P-loop) Lys562 (amino acid numbering is as per the Axl RTK)- Glu585 (α C-helix) represented “active state” of RTKs where the dissociation of these interactions is described as “inactive state”. During these RTKs activation process, the P-loop and α C-helix outward (inactive) or inward (active) rotation with help activation loop folded into β -sheet in active state, the activation loop refolded into helical structure exists as inactive state. This can be confirmed with the help four consecutive hydrophobic residues movement from N-terminus to C-terminus as called regulatory spine. This can be generated as four non-consecutive residue motifs from four kinase domains as P-loop (Lys), α C-helix (Glu), DFG motif (Phe), catalytic loop (His). there are eight non-consecutive residues from N-terminal to C-terminal represented as catalytic spine. The catalytic spine is ATP bound region for catalytic activity of tyrosine kinases as the two non-contiguous motifs of these spines are connected with catalytic loop and α F-helix region in the C-terminal domain of tyrosine kinases (Robinson, 2013; Mohanty et al., 2016; Hu et al., 2015).

The kinase activity of RTK is triggered by its binding to ATP and Mg^{2+} that results in the transfer of γ -phosphate group to tyrosine containing protein target. The ATP binding cleft is located between the N- and C-terminal lobes, and at the hinge region connecting the two lobes (Kornev et al., 2006). The structurally important regions required for the activity of a protein tyrosine kinase include, the P-loop located between β 1 and β 2 strands, catalytic helix (α -helix) in the N-terminal lobe comprising the essential amino acid (Glu585, Axl amino acid numbering) with its side chain fluctuating between the active and inactive states of kinase. The distinction between the active and inactive states is also based upon the α -helical movement towards or away from the ATP binding site. The presence of catalytically important Lys567 (close to P-loop) - Glu585 (α -helix) is an essential ionic interaction in the active Axl kinase from the crystal structure in protein data bank (PDB). The disordered activation loop (689-724) in the C-terminal lobe has altered conformational states that are variable among the kinase structures reported so far. An ionic interaction between the side chains of Asp581 (α C-helix) and Lys695 (activation loop) is important in the kinase structure and allostery. The synchronous fluctuations in the P-loop, α -helix and activation loop leads to spatial alteration in the shape of the enzyme active site pocket and distinct structural

features such as the inward/outward rotation of α -helix and expansion of the activation loop. The Lys567 - Glu585 salt bridge is the indication for the active state of TAM RTKs. A kinase domain has two kinds of active sites; regulatory substrate site and catalytic active site that become available during allosteric competitive inhibitor binding pathways in the cellular signal transduction process. Structure analyses revealed the presence of two non-contiguous structural motifs, regulatory and catalytic spines (Robinson, 2013, Mohanty et al., 2016, Hu et al., 2015) that are required for stabilizing the protein in the active state.

1.2. Severe acute respiratory syndrome coronavirus 2

The severe acute respiratory syndrome coronavirus-2 (SARS CoV-2) has caused a global pandemic of the coronavirus disease 2019 (COVID-19) during the last 40 months. COVID-19 has been the single major cause of death due to any disease within a short span of time. SARS CoV-2 was first reported in individuals known to have been in contact with wildlife animals at the live animal and seafood market in Jiangnan District, Wuhan (Zhu et al., 2020). SARS CoV-2 is similar to SARS CoV (2003 to 2005), Middle East respiratory syndrome coronavirus (MERS CoV) (2012 to 2013) and other human CoVs in the 20th century that has led to epidemics resulting in severe respiratory diseases and deaths (Guruprasad, 2021a). These viruses harbour ~ 30K bp single stranded positive-sense RNA genome. SARS CoVs enter human cells through fusion of viral and host cellular membranes mediated by the interaction between viral spike protein and human angiotensin converting enzyme-2 (ACE-2) (Guruprasad, 2020; 2021b; Li et al., 2003; Shang et al., 2020). The SARS CoV-2 spike protein is a heavily glycosylated homo-trimeric protein with ~1,273 amino acids and the sequence region (amino acids 333-520) constitutes the receptor-binding domain (RBD) that interacts with human ACE-2 receptor. The three dimensional structures of the spike protein apo and RBD bound forms to human ACE-2 receptor are available in the public domain (Wang et al., 2020; Xiao et al., 2021; Xu et al., 2021). Viruses acquire mutations over a period of time during host infection giving rise to new sequence variants. RNA viruses have much higher mutation rates compared to DNA viruses. The viruses that undergo favourable mutations continue to persist in host. Due to mutations, the viruses might gain ability to evade detection by specific viral diagnostic tests, or decreased susceptibility to therapeutic agents, such as, monoclonal antibodies and small molecule drugs. Some mutations can produce viruses with new antigenic determinants and the antigenically altered viruses may be able to cause disease in previously resistant or immune hosts or cause vaccine rejection (Fleischmann, 1996). The wild-type refers to first reported strain of the human SARS CoV-2 virus isolated from patient

in Wuhan, China (NCBI_id: NC_045512) (Wu et al., 2020). With respect to wild-type proteins, the mutations and sequence variants are collected through complete genome sequence and epidemiological studies of SARS CoV-2 strains across populations from various geographical locations and different times. These sequences have been deposited in the NCBI (www.ncbi.nlm.nih.gov) and GISAID (<https://www.gisaid.org/>) databases. Sequence analyses have reported deletions, insertions and substitution mutations in all SARS CoV-2 proteins including the spike protein (Guruprasad, 2021b; Mohammadi et al., 2021) demonstrating that SARS CoV-2 has an innate ability to undergo mutations rapidly. SARS CoV-2 vaccines show protective efficacy towards humans by providing neutralizing antibodies which recognize the viral spike protein (Kyriakidis et al., 2021). The effects of spike protein mutations in SARS CoV-2 on the neutralization of antibodies have been studied (Rees-Spear et al., 2021). Genome sequencing and protein sequence analyses have shown the emergence and persistence of some SARS CoV-2 spike protein mutations in subsequent generations during human infection. One of the early identified mutations in spike protein, D614G is associated with lower RT-PCR cycle thresholds suggestive of higher viral loads but not with increased disease severity (Korber et al., 2020). The D614G mutant spike protein increases SARS CoV-2 infection of multiple human cell types compared to the wild-type strain (Daniloski et al., 2021) and efficiency of viral entry with enhanced ACE-2 binding affinity (Ozono et al., 2021) by assembling more functional spike protein into the virion (Zhang et al., 2020). Epidemiological evidence suggests that the D614G variant has increased ability to spread more quickly than viruses without this mutation. Therefore, the D614G mutant has become dominant in the SARS CoV-2 spike protein. The Phylogenetic Assignment of Named Global Outbreak Lineages (PANGOLIN) software tool (Rambaut et al., 2020) implements a dynamic and rational nomenclature of the SARS CoV-2 strains. The PANGO lineages available at website (<https://www.cov-lineages.org>) refers to the cluster of sequences associated with distinct geographical locations with evidence of onward spread and captures the emerging trends of mutations from genomic epidemiological surveillance and outbreak investigations. The US government SARS CoV-2 Interagency Group (SIG), initially developed a variant classification scheme that defines three classes of SARS CoV-2 variants; variant of interest, variant of concern and variant of high consequence. A variant of interest has specific genetic markers associated with changes in receptor binding, increased disease severity, reduced efficacy of treatments, reduced neutralization by antibodies generated against previous infection or vaccination, potential

diagnostic impact, increased proportion of clusters of cases and therefore increased transmissibility.

A variant of concern must display evidence of increased transmissibility, more severe disease with evidence of reduced effectiveness of treatments or vaccines, or failure in diagnosis detection, treatments, or vaccines leading to increased hospitalizations or deaths. The SARS CoV-2 variants are classified into four classes; variant being monitored, variant of interest, variant of concern and variant of high consequence. The lineages for variant being monitored and variant of concern along with their WHO label and mutations in the spike protein are documented. Currently, there are no mutations classified under variant of interest or variant of high consequence. The B.1.1.7 lineage first detected in UK during September 2020 was subsequently reported in several countries including India. The substitution mutations in the RBD of spike protein in this lineage were; E484K, S494P, N501Y. This lineage is attributed to 50% increased transmissibility (Davies et al., 2021), increased severity based on hospitalizations and fatality rates compared to other variants (Horby et al., 2021). The B.1.351 lineage was first identified in South Africa during early October 2020, Zambia during late December 2020 and subsequently reported in several countries. The substitution mutations in the RBD of spike protein in this lineage were; K417N, E484K, N501Y. This lineage is also attributed to 50% increased transmissibility (Patone et al., 2021). The E484K mutation may affect neutralization by some polyclonal and monoclonal antibodies (Patone et al., 2021). The P.1 variant was first reported in Japan in travelers from Brazil, subsequently, Manaus, in the Amazon region and also in the United States at the end of January 2021. The P.1 lineage mutations in the RBD of spike protein were; K417T, E484K, N501Y. The P.1 variant mutations may affect its transmissibility and antigenic profile resulting in its decreased ability to recognize antibodies generated through a previous viral natural infection or through vaccination (Harvey et al., 2021). The B.1.617.2 lineage contains the mutations; K417N, L452R, T478K in the RBD of spike protein. The B.1.427 and B.1.429 variants of SARS CoV-2 were first detected in southern California, USA during early 2021 that comprises the L452R mutation associated with the RBD. Both variants have 20% increased transmissibility (Deng et al., 2021). The B.1.525 lineage comprises the E484K mutation and the B.1.526 lineage comprises mutations; L452R, S477N, E484K in the spike protein RBD. The sub-lineages of B.1.617; B.1.617.1 comprises the mutations; L452R, E484Q; B.1.617.3 comprises mutations; L452R, E484Q and B.1.672.2 (delta variant according to WHO classification) and variant of concern comprises mutations; K417N, L452R, T478K. The P.2 lineage first reported in Brazil during October 2020 has four mutations in the

entire spike protein with the E484K mutation in RBD. The B.1.621 lineage that originated in Colombia during January 2021 has the mutations; E484K and N501Y in RBD.

1.3. Materials and methods

1.3.1. Protein sequence databases

BLAST: The Basic Local Alignment Search Tool (BLAST) is a rapid basic nucleotide and protein sequence alignment tool from sequence databases by sequence searches to generation of phylogenetic tree of various organisms. The DNA and RNA, protein sequence datasets are deposited in many online libraries (NCBI, UNIPROT online databases). BLAST is a sequence alignment search tool through web resources to connect different species of sequence data in a stochastic and robust manner. It searches the different genetic sequences in multiple contexts as specific sequence motif and its identification search, RNA and DNA, protein database searches of respective gene location and mapping of similar gene sequences (Altschul et al., 1990). This algorithm has a large pool of resource to sort the protein sequences of desired results. The sequence alignments from BLAST can be a guide to find a homologous protein of known structure that could be used as a template for homology modeling. Based upon this sequence alignment, the Phyre2 server proposed theoretical protein homology models to compare with actual sequence based designed models (Shen, & Sali, 2006; Laskowski et al., 1993).

1.3.2. Multiple sequence alignment of protein sequences

Multiple sequence alignment is a series of stacked algorithms resulting in the evolutionary hierarchy alignment of genomic sequences in systematic manner from one or more organisms. This alignment series describes about the evolutionary events like genetic insertions, mutations, deletions and overall rearrangement, that is further modified in younger evolutionary generations. These alterations are occurred in DNA, RNA and protein sequences. The comparative sequence analysis can be done with BLAST and online multiple sequence alignment webserver by genetic sequence databases only. The multiple sequence alignment is a fundamental technique to align the genomic sequences in the modelling methods of interdisciplinary biological sciences referred to as Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (Sievers, & Higgins, 2018) in scientific community. It is one of the top most cited research papers in fundamental biological research field (Chatzou et al., 2016). The multiple sequence alignment is followed by phylogenetic tree generations and domain analysis and their reconstruction of selective motif from sequence information. Major purpose of multiple sequence alignment is to align the structural genetic

sequences into a way such that it will represent evolutionary, structural and functional relationship among different organisms. For building a multiple sequence alignment there is a need for scoring functions to quantitative merits in a relative manner with respect to sequence relationships. It can be estimated by using best scoring model function with respect to sequence data (Madeira et al., 2022).

1.3.3. Protein homology modelling and validation

The hierarchy of protein structural information is defined at four levels; primary, secondary (α -helical and β -sheet), tertiary and quaternary structures. The functional aspect of a protein structure is represented as three dimensional spatial arrangement of polypeptide chains consisting of main chain and side chains (Luthy et al., 1992). The protein structures are solved using X-ray protein crystallography, nuclear magnetic resonance (NMR) and cryo-electron microscopy in three dimensional form as stored in PDB format in many webserver (RCSB PDB.org) (Berman et al., 2007). The three dimensional structure of a protein provides useful insights of morphological and functional activity at molecular basis of protein conformers. Because in the entire protista, animalia and plantae kingdoms 1/3 of whole proteome sequences are related to one of the protein sequences of known structure. As there are inherent difficulties in solving the structures of proteins, the building of protein three dimensional model structures has gained popularity as it is a reliable method. Based upon this criteria, the rate of modelled protein structures by comparative structure modelling method is a far higher order than known experimentally determined structures. The comparative modelling methods have high accuracy for low resolution X-ray structure, which are very hard to render through single X-ray diffraction method. The number of methods to design protein models using homology modelling methods are increasing rapidly to evolving change in computational biophysics research area. The protein homology modeling technique is computationally viable protein structural prediction method to deduce the three dimensional protein model as replica of existing template in tertiary structure and fill any the missing residues in secondary structural or loop regions and from the given protein sequences (Cavasotto & Phatak, 2009). It is one of the efficient tools to generate the non-existent and difficult to crystallize proteins with the existing laboratory techniques. The biological simulations results have been depending upon the maximum quality of three dimensional homology models. The homology modelling is key computational technique to enhance the protein ligand and protein-protein, protein-DNA/RNA interactions in biological simulations with statistical mechanics (Harding et al., 2002). Homology modelling of protein is very crucial in mutations induced in wild type protein sequences and their

mutated homology model generation is important for mutation analysis of these models in docking and MD simulated systems (Muhammed, Aki-Yalcin., 2019). The virtual model generation of three dimensional protein structures are essential to search for effective inhibitor of particular protein active site models in drug discovery process. It would enhance the drug discovery possible with practical unknown and inaccessible protein models in experimental process. The protein three dimensional structural interactions with inhibitor and substrate molecules can be possible with these virtual model generations without physical need of protein samples. There are key steps involving to generate the effective comparative modelling of protein from genetic sequences. To identify selective homology gene sequence extracted from various genetic sequence libraries (NCBI database) with the help of phylogenic genetic sequence order. Based upon the multiple sequence alignment of desired template and unknown protein sequence or known target protein superpose to identify the % of sequence identity to generate maximum three dimensional protein structural model accuracy (Kopp & Schwede, 2004). The model building and loop refinement, side chain rotamers modeling is done with model build python script from Modeller (Shen, & Sali, 2016; Yang et al., 2012) or SWISS-MODEL (Bordoli et al., 2009). The selective lower DOPE scores models are further validated with model optimization. The model quality evaluation of given protein sequence to generate the replica of structural morphology from template structure. The tertiary homology model generation methods include Modeller, I-TASSER, AlphaFoldDB, SWISS-MODEL, Molecule operating Environment (MOE), PHYRE2, PRIME webtools to build the protein three dimensional models from genetic sequences (Varadi et al., 2022; Waterhouse et al., 2018; Kelley et al., 2015). The protein model validation is done by stereochemical model validation for allowed residues region with Ramachandran plots and with ERRAT, PROCHECK validate three dimensional spatial orientation of homology models (Laskowski et al., 1993; Colovos, Yeates., 1993; Ramachandran et al., 1963).

1.3.4. Protein data bank

Protein data bank (PDB) is a web-based protein structure database available worldwide from 1999 onwards. It is a single archive of protein structural database sometimes in complex with various biomolecules from structural biologists. It is a web based server (wwPDB) where the archive datasets are obtained from structural biologists, biochemists, protein crystallographers who solved the protein structures using NMR, cryo-electron microscopy and X-ray crystallography. This was organized by four members of institutions as PDB in Europe (PDBe), Biological Magnetic Resonance Data Bank (BMRB) and Research Collaboratory for Structural Bioinformatics (RCSB)

PDB from USA and PDB bank from Japan (PDBj) (Tong, Ranganathan, 2013). These entities collectively maintain the RCSB PDB (<https://www.rcsb.org/>) web archive of single protein database of any missing structural biomolecular data from global scientific community. It is an accessible web archive server as free of cost to provide experimental and computed protein structural models included from AlphaFold and Model Archive databases (Bittrich et al., 2022). From 2020 onwards, there is enormous increase in the deposited crystal and computational protein structures into these databases due to COVID19 pandemic. Therefore, the PDB database hierarchy has been done in PDB entry of single crystal protein structure, chemical entity of small molecules, multiple instances of any individual chemical entity, assembly of all polymorphic chains and other non-polymeric chemical entity considered assembly with unique PDB id in RCSB PDB databases (Burley et al., 2022). This databank provides high volume of protein sequence and three dimensional coordinate system of individual protein structures including cryo-electron microscopy structures at ultra-low temperature crystallography with lasers. As on 27th July, 2023, there are 207,791 experimental structures from the PDB and 1,068,577 computed structures available at www.rcsb.org/.

1.3.4.1. Small molecule databases

In ligand-based and structure-based drug design methods, the small organic molecules three dimensional coordinates are extracted from various small molecule drug databases. The small molecule crystal information files (cif) are deposited from large number of scientists including multiple domains scientific community groups around the globe. Drug hunter database provides (<https://drughunter.com/molecules-of-the-month/>) all the FDA approved and commercially available drugs in market with a monthly update. The national cancer institution is one of the large cancer chemotherapeutics database (https://ntp.niehs.nih.gov/databases_tools/). There are many small molecule databases generated from a pool of chemical libraries such as ChEMBL32, MolProt, PubChem, Zinc, ChemDiv and ChemSpace (Bates et al., 1995; Hollingshead et al., 2022). These libraries are database repositories of small molecular inhibitors of millions of compounds for selective protein active site target. The pharmacokinetics studies of small molecules are done with SwissADME webserver tools. These studies are validated with known virtual and experimental adsorption, distribution, metabolism, excretion and toxicology (ADMET) (Potts et al., 1992; Daina et al., 2017) results validated with phase II and III clinical trials of small molecules for potency towards a specified disease. The pharmacokinetic properties of all screened drugs are described with SwissADME, a web based online server tool.

1.3.4.2. Pharmacophore generation and virtual screening of chemical libraries

The computer aided drug discovery has major bottlenecks in the design of suitable drug molecules for specific protein active site target with different and viable chemical environment entities with unique molecular mass such that they qualify as drug molecules. In this direction, one of the major challenges is to select a specific number of small molecular hits to satisfy the conditions of approval for drug-like candidates by agencies such as FDA. Therefore, a modern computational advancement is by effective retrieval of pharmacophore models based drug-like hit molecular query to initiate the hunt for target protein binders towards a disease condition with the help of online search databases; Pharmit (Sunseri, Koes, 2016), iDrug (Chen, & Cheng, Li, 2020), Dock Blaster (Irwin et al., 2009) and related methods (Gan et al., 2023; Singh, & Chaput, Villoutreix, 2021). The virtual screening can be done based upon morphology of protein active site and the chemical environment around the reference small molecular inhibitor (pharmacophores). However, the virtual screening of small molecules is very prominent for drug discovery in the modern data era due to the availability of advanced algorithms and high performance computational facilities around the world (Folding@Home, supercomputing facilities availability in Switzerland, Germany USA, UK, France, China and Japan) (Maia et al., 2020). The virtual screening can be easily done with high throughput screening with a large server of computer cluster to decipher the desired small molecular entities which is cost effective and has greater reliability in finding appropriate pharmaceutically active relevant hit molecules from large virtual databases. Virtual screening is an *in silico* method for structure based query against protein three dimensional structure active site to select desired hit molecules by searching various small molecules data libraries. Virtual screening works like a funnel to segregate the small molecules based upon their pharmacophore features and molecular mass and should be viable to inhibit cavities in the environment of protein binding sites. This technique has numerous advantages during the drug development and optimization to search for bioactive hit molecules. These virtual databases comprise molecules with appropriate biological activity to optimize a candidate molecule of interest for required protein targets to finally lead to a drug molecule. After selection of optimized hit molecules, the pharmacodynamics and toxicological data prediction can be deduced from their respective structural and physiological properties (Lill, 2013; Daina, Michielin, Zoete. 2017). The virtual screening was first developed in 1980 due to the computational advancement and abundance of synthetic biochemical active molecules to target many virtual three dimensional protein coordinates available in PDB databases (Anderson, 2003). While the experimental *in-vitro*

screening of active pharmaceutical ingredients (API) is more tedious and highly expensive in terms of traditional drug discovery like high throughput screening, the virtual screening is an excellent in silico tool to find the most probable and appropriate drug-like entities to bind the selective and precise active site regions of particular protein targets. This methodology is cost efficient and without wastage of external solvent chemicals and without the requirement of any other dedicated instrumentation (Lill, 2013).

The virtual screening is a systematic computer aided drug discovery tool to design the effective small molecule inhibitors based upon pharmacophore features of known drug chemical entities and protein active site complementarity with effective non-bonding interactions mediated with the surrounded residues. The selection of pharmacophore features is based on the complementarity of protein receptor – ligand or inhibitor complex. ZINCPharmer (Koes, & Camacho, 2012), Pharmit (Sunseri & Koes, 2016) are some of the popular pharmacophore generation online webserver. The pharmacophore generating methods are used to generate the best pharmacophore for the design of hit molecules binding to selective target protein. The best generated pharmacophore is used to screen online small molecule databases. The databases available in the Pharmit server are used for pharmacophore based virtual screening (Sunseri & Koes, 2016). It screened libraries of millions of compounds based upon pharmacophore and molecular shape modalities with advanced pre-built structural screening algorithms through the online webserver. The possible features available in the Pharmit server are, hydrogen bond donor, hydrogen bond acceptor, hydrophobic and ring aromatic. Finally, it provides comprehensive query optimized hit molecules with online platform for structure based virtual screening method. The virtual screening is done based upon selective protein active site domain and ligand pharmacophore chemical entity surrounded by residue environment in protein of interest complexed with known inhibitors. The molecules obtained based on lower root mean square deviation (RMSD) and fewer rotatable bonds are often selected. An in-house database of the downloaded molecules is prepared after adding hydrogens and their structure refinement using structure visualization tools such as Discovery Studio 3.5. The identified molecules are passed through the virtual screening protocol using PyRx (Dallakyan et al., 2015) by docking the molecules into the binding site of Ligand. The screened-in molecules above a certain binding energy are selected for molecular docking studies.

1.3.5. Computer aided drug design

The structure based drug design has gained popularity in the last 25 years due to the availability of enhanced computational genomic data and proteomic approaches, molecular modeling

methodologies to design effective protein active site inhibitors. The structure based inhibitor design has many high hierarchical levels to select the protein target and detect the model active site for selective inhibitor design. This method is continuous and iterative cyclic process until the lead optimized hit molecules entered in phase-I clinical trials often starting from a chemical scaffold. The initial phase cycle of structure based drug design has selective gene cloning, protein purification and experimental techniques for structure determination using instrumental techniques such as X-ray crystallography, nuclear magnetic resonance or cryo-electron microscopy. Alternatively, homology modeling of protein structures or structures based on distant relationships can be generated. The selective active sites or allosteric sites can be defined from protein structures. Here the computational algorithms have a key role to identify most probable binding sites to further validate them as selective drug binding sites (Mokaya et al., 2023). Some selective molecules to bind the active sites can be designed based on pharmacophore features of substrate or inhibitor molecules by employing computational methods.

The molecular docking is carried out in an iterative manner so that the lead score hit molecules are constantly modified for selective drug target bound at receptor active site with highest binding affinity. The top priority selective hit molecules are synthesized or commercially purchased with highest purity to validate the receptor-hit molecules using experimental binding assays. Various in-vitro, in-vivo and cell toxicology studies are required to validate the biological mimic of hit molecules from the biochemical studies. The drug target and lead optimized molecules must have good appraisal results for bioavailability and lead molecules affinity towards target protein in general should be good. Then the modified lead molecule moves from micromolar inhibition zone to nanomolar inhibition zone during assays in various biological systems. If all the laboratory tests are positive, the lead molecules enter into various phases of clinical trials along with pharmacology studies. In computational methods, the targeted receptors can mimic the experimental conditions with the known selective number of lead hit molecules binding to facilitate the more successful in-vitro and in-vivo studies by providing biomimicry studies for lead target of receptor-hit molecules complexes (Anderson, 2003).

1.3.5.1. FDA approved kinase inhibitors and their receptor targets

(Leone et al., 2023; Lee, & Yeoh, Low, 2023; Roskoski, 2022; Roskoski, 2023; Lui et al., 2022;

S.No	Target kinase	FDA approved inhibitors (year)
01	ALK	Alectinib (2015), Brigatinib (2017), Crizotinib (2011), Ceritinib (2014) , Lorlatinib (2018),
02	BCR-ABL	Asciminib (2021), Bosutinib (2012), Dasatinib (2006), Imatinib (2001), Nilotinib (2007), Ponatinib (2012),
03	BTK	Acalabrutinib (2017), Ibrutinib (2013), Zanubrutinib (2019)
04	CDK4/6	Abemaciclib (2017), Palbociclib (2015), Ribociclib (2017), Trilaciclib (2021),
05	ErbB2	Neratinib (2017), Tucatinib (2020)
	EGFR	Afatinib (2013), Brigatinib (2017), Dacomitinib (2018), Erlotinib (2004), Gefitinib (2003), Lapatinib (2007), Mobocertinib(2021), Neratinib (2017), Olmutinib (2016), Osimertinib (2017), Osimertinib (2015)
06	Flt3	Gilteritinib (2018), Midostaurin (2017),
07	JAK1/2/3/STAT	Abrocitinib (2022), Baricitinib (2018), Fedratinib (2019), Pacritinib (2022), Ruxolitinib (2011), Tofacitinib (2012), Upadacitinib (2019),
08	Kit	Ripretinib (2020),
09	MEK1/2	Binimetinib (2018), Cobimetinib (2015), Selumetinib (2020), Trametinib (2013)
10	MET	Capmatinib (2020), Tepotinib (2021),
11	mTOR	Everolimus (2009), Sirolimus (1999)
12	PI3K	Copanlisib (2017), Temsirolimus (2007),
13	ROS1 (TAM)	Crizotinib (2011), Entrectinib (2019) Cabozantinib (2012), Erdafitinib (2019)
14	RAF	Dabrafenib (2013), Vemurafenib (2011), Encorafenib (2018)
15	RET	Selpercatinib (2020), Cabozantinib (2012), Pralsetinib (2020), Selpercatinib (2020),
16	ROCK1/2	Netarsudil (2018)
17	Syk	Fostamatinib (2018), R406 (Fostamatinib) (2018),
18	TRK	Larotrectinib (2018), Ibrutinib (2013), Entrectinib (2019), Acalabrutinib (2017), Zanubrutinib (2019),Deucravacitinib (2022), Pexidartinib (2019), Larotrectinib (2018),
19	VEGFR2	Sunitinib (2006), Erdafitinib (2019), Axitinib (2012),Cabozantinib (2012), Futibatinib (2022), Infigratinib (2021),Lenvatinib (2015), Nintedanib (2014), Pazopanib (2009), Pemigatinib (2020), Regorafenib (2012), Sorafenib (2005), Tivozanib (2021), Vandetanib (2011),

1.3.6. Molecular docking

Molecular docking is one of the *in silico* drug discovery tools to obtain a first crude approximation about the ligand orientation into receptor active site or a probable pose of host-guest complex. In the case of protein and ligand or inhibitor complex, molecular docking is the key tool to investigate the best docked pose of a ligand molecule in the enzyme active site. Molecular docking can be performed in multiple ways, one is blind docking and the other is selective active pocket region of protein for docking location. If the active pocket of protein is unknown, the blind docking is the initial optimization in docking process later it should proceed into selective active site region. It is an iterative process and multiple self-consistent procedures would give better docking pose for a given protein-ligand complex. The aim of docking procedure, is to search the effective and predictive active drug pose to bind the receptor as principal binding space occupied in three dimensional protein structure. The docking is an *in silico* plotting ligand pose optimized for all possible three dimensional ligand conformers in possible chemical space of binding location in three dimensional protein structure coordinates. Generally, the docking procedure depends mainly upon two factors as effective docking orientation in three dimensional space and docking scoring functions to achieve maximum success for docking pose for three dimensional small molecules. The docking methods are classified based upon the chemical environment around the hit molecules and the flexibility of the receptor or protein active site pose in three dimensional structure. The hit molecules are quite flexible in docking pose due the higher chance to orient the hit molecule in active site of protein and chemically rigid manner to interact with active residues of protein. Therefore the flexible docking process is involving both protein and ligand chemical space to build non-bonding interactions among them. However, the flexible docking is more computationally intensive and efficient in terms of accuracy and probability to find ideal hit molecule's docking poses. In general, the protein is in a rigid manner, the ligand or inhibitor molecules are quite flexible and dynamic in active docking pose. Here some of the docking methods follow the above procedure to fit the hit molecules in active site of protein such as AutoDock4, CDOCKER, PyRx, LibDock and etc (Kramer et al., 1999; Wu, 2003; Morris et al., 2009; Torres & Olson., 2019). But GOLD, DOCK, Glide docking procedure predicted the docking pose based upon docking scoring function as both receptor and ligand are flexible and complementary to each other (Friesner et al., 2004; Lang et al., 2009; Verdonk et al., 2003; Zhao & Caflisch, 2013). However, the both ligand and protein molecules are rigid positions in ZDock and RDCOK methods (Chen et al., 2003; Li et al., 2003). In molecular docking technique, various non-bonding interactions involved between

ligand and protein include electrostatic attraction forces, hydrogen bonds, hydrophilic and hydrophobic interactions of residues and partial charges on specific functional group from residues in vicinity on the tertiary structures. The docking results are compared with initial reference crystal PDB structure to consider for reference to obtain good docking poses of ligands. The sorting of docking poses are based upon maximum non-bonding interactions and suitable docking binding free energies and best docking scorings for efficient docking site search method. The molecular docking can be done in a systematic procedure in AutoDock; i) to view protein and hit molecules in three dimensions, rotate and scale in real time. ii) Add all hydrogens or just non-polar hydrogens to protein iii) Assign partial atomic charges to the ligand and the macromolecule (Gasteiger or Kollman United Atom charges). iv) Merge non-polar hydrogens and their charges with their parent carbon atom. v) Set up rotatable bonds in the ligand using a graphical version of AutoDock4 tools or any graphical dock viewer. vi) Set up the AutoGrid Parameter File (GPF) using a visual representation of the grid box and slider-based widgets. vii) Set up the AutoDock Parameter File using AutoDock4 module. viii) Launch AutoGrid and AutoDock with parameter file for calculating the grid of the molecular basis and flexible docking calculations done for ligand to obtained optimum docking pose in the presence of active site of protein. ix) Read the results of an AutoDock job and graphically display them, while comparing with crystal structure of protein. x) View iso-contoured AutoGrid affinity maps and binding free energies (docking scores obtained from Glide, GOLD docking method).

1.3.7. Absorption, distribution, metabolism, excretion and toxicology studies

The assessment of ADMET properties are required to find the drug-like properties of bioactive molecules, potent enough to reach the target site located in various body parts and that the chemical entity stays just long enough to perform its biological activities. The ADMET properties are the pharmacokinetics properties, effective clinical cell toxicity and clinical interferences caused in normal body physiological processes after clinical drug administration in healthy individuals or patients. These are the properties of clinical drug molecules to maintain optimal pharmacophore drug like properties while drug action at the desired site in our body. These properties are very crucial to understand the experimental in-vivo, in-vitro pharmacological properties with selective drug studies. The pharmacokinetic properties of the virtual screened molecules were assessed with SwissADME (<http://www.swissadme.ch>) (Daina, Michielin, & Zoete., 2017) web-based server tool. The drug likeness properties are represented in the form of topological polar surface area (TPSA), Consensus Log $P_{o/w}$, LogS (ESOL), LogKp (skin permeation) (cm/s), synthetic

accessibility (Daina, Michielin, & Zoete., 2014, Ertl, Rohde, & Selzer., 2000, Potts, & Guy., 1992). ADMET are the key notable pharmacokinetic properties to qualify the drug like small molecules for clinically approved pharmacological properties.

The Lipinski's rule of Five (Ro5) were developed in Pfizer by Christopher Lipinski (Lipinski et al., 2004) to validate small organic molecules for drug-likeness for oral administration. Most of the small molecular inhibitors follows these rules to act as drug like candidates in oral route. In general, the Lipinski's rules suggest the small molecules has molar mass less than 500 g per mol. The ligand has maximum five hydrogen bond donors and not more than ten hydrogen bond acceptors and the body intake water permeability (partition functions $\text{Log } P > 5$) should be high value (Mahgoub, Atatreh, & Ghattas, 2022).

1.3.8. Molecular dynamics simulations

The first three dimensional enzyme structure is characterized by X-ray crystallography (Phillips et al., 1966.). Further analysis is carried out to emphasize the importance of protein flexible conformational changes while binding with ligand or substrate in its active site. However, during the early development of computers (Levitt & Lifson, 1969), the first MD simulations was done more than 60 years ago. During this timelines, the protein-ligand and genomic sequence analysis were very limited due to the lack high speed computational power (McCammon et al., 1977). But, in the 21st century of the digital era, the fast forward moving world with high speed computational enhancement, these protein-ligand interactions and higher organism genomic/proteomic analysis is quite possible. Nevertheless, the fundamental understanding of the protein conformer and its dynamics when complexed with suitable drug or hit molecules could be key components in drug discovery to further computationally validate in silico drug discovery methodology. In recent years, the computational drug discovery received prominence due to theoretical protein structural validation and virtual screening of small molecules with the help of cost effective high performance computation facility (Shaw et al., 2007; Smith et al., 2002). Therefore the molecular mechanics of protein drug complexes with longer range MD simulations is studied to generate possible hit molecules validated for a desired protein target. The MD simulations is one of computational techniques to effectively simulate the molecular motions from host-guest interactions in three dimensional space. It is based upon of numerical solutions of classical Newtonian mechanics (Dahiyat & Mayo, 1997; John & Sali, 2003). The molecular simulations of individual atomic velocities, positions, accelerated dihedrals of bonding and non-bonding motion equations from protein-ligand molecular systems generated the molecular trajectory data with

respect to simulations time. These molecular equations provide the total and potential, kinetic energies of bonded atoms and van der Waals energy and Lennard-Jones potentials of non-bonded interactions, charged or electrostatic and columbic energies (Childers & Daggett, 2017; Geng et al., 2019) are also calculated with the help of numerical solutions in terms of Newtonian mechanics (Pace et al., 1996). The digital modern computational facilities have expanded by including high performance computing power in the recent times. Therefore, MD simulations got huge popularity in drug discovery to integrate with healthcare for looking into the solutions of precise medicine validation without intervention of in-vivo drug studies in the initial stages. It can save a lot of time and experimental flaws with tedious high throughput in-vitro drug discovery process. MD simulations provide the detailed information about apo protein, protein – protein and protein – ligand interactions in water sphere of statistical dynamics perspective. Various modules describe important molecular information regarding protein-ligand interactions from MD simulations. Several MD simulations packages available such as CHARMM (Jo et al., 2008), AMBER (Case et al., 2005), LAMMPS (Grindon et al., 2004), NAMD (Nelson et al., 1996), and GROMACS (van Gunsteren & Berendsen, 1987), Desmond (Bowers et al., 2006) are available for simulating various types of computational simulation problems in different domains of science and technology. The force - mass equation in the second Newtonian law is used to calculate the relative force constant values on dynamic motions for overall molecular systems studied in MD simulations.

1.3.8.1. Force fields

Force fields is collection of mathematical expressions from molecular geometry energy represented equations of Cartesian coordinates of atoms in a molecular structure. The molecular geometry can be any bio-macromolecule or small organic molecule or any chemical entity represented in three dimensional space including higher order material frame work. It describes the total energy of the entire molecular system in three dimensional structural space. In general a force field is represented into two terms. First term is represented as molecular bonded interactions including covalent and electrostatic bond lengths, bond angles and torsional angles. The second term is represented as non-bonding molecular interactions including the van der Waals, dipoles, induced dipoles, London forces, charge particle attraction forces measured by Lennard-Jones potentials, Coulomb's law mathematical expressions (Dahiyat & Mayo, 1997; John & Sali, 2003). This mathematical equation represents the entire force fields bonding and non-boning of any molecular system with surrounding molecules.

$$E_{\text{total}} = \sum_{\text{bonds}} N_i^{\text{bond}} (r_i - r_{\text{eq}})^2 + \sum_{\text{Angles}} N_i^{\text{angle}} (\theta_i - \theta_{\text{eq}})^2 + \sum_{\text{dihedral}} \frac{V_n}{2} [1 + \cos(n_i \phi_i - \delta_i)]$$

$$+ \sum_{i > j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}^{12}} \right] + \sum_{\text{non-bonding}} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right]$$

These parameters can be obtained from quantum chemical calculations such as ab-initio and semi-empirical methods and the experimental parameters such as X-ray diffraction, NMR and Raman spectroscopic data etc (Weiner & Kollman, 1981). In a molecular system, the molecular bonds act like springs tied together in an atomic system to under giggling, vibrational states occurring during molecular oscillations by external force or energy. The force fields are defined as overall molecular, electronic, vibrational, rotational, translational energy terms expressed for a broad range of macromolecules or small organic molecules represented in three dimensional space. In most cases the force fields should be quite simple to analyse the small molecular system. The three dimensional building blocks of these tiny unit cells replicate into true crystal structure or total macromolecules approximated classical total energy can be extracted from MD simulations trajectory data for biomimicry of real cell based properties with the presence of water solvation environments. The classical thermodynamic based parameters are majorly included to represent the force fields for any molecular systems from periodic table of elements (Rappe et al., 1992). The most popular molecular mechanics force fields from different force field development groups are AMBER (Malolepsza et al., 2010), CHARMM (Brooks et al., 2009), GROMOS, OPLS (Jorgensen et al., 1996) and COMPASS (Sun et al., 1998). Different versions of force fields are employed to achieve higher computational accuracy and efficacy for various macro and small molecules [CHARMM19, CHARMM22, CHARMM27, AMBER91, AMBER94, AMBER96, AMBER99SB, AMBER14SB, GROMOS96, GROMOS45A3 and GROMOS53A5]. Generally these force fields are compatible with both materials and biomolecules to run MD simulations in higher order or long range timescales. But some packages are more specific to certain systems; AMBER – Carbohydrates, bio-molecules DNA, RNA, Protein, lipid-bilayer; CHARMM and

CVFF (Sun et al., 2016; Lin, & MacKerell, 2019) are mostly compatible with different types of materials in various crystal systems.

1.3.8.2. H-atoms addition to apo form with H++ server

In general, the DNA/RNA have an overall negative formal charge, but the proteins have different charges associated with varying isoelectric points on the basis of various amino acid sequences (primary structure) that make up to their secondary and tertiary structures. In GROMACS and AMBER MD simulations, these small atoms of hydrogen play huge role to predict their overall formal charge and major protonation charges varies in certain amino acids like histidine, aspartate, tyrosine, lysine, arginine, glutamic acid. These types of amino acids have an extra amine and carboxylic group which can be in protonated or deprotonated states in a huge pool of hydrosphere in MD simulations (Anandakrishnan, Aguilar, & Onufriev, 2012). Before submitting for energy minimization of molecules and ions addition to neutralize the charge, the protein conformer should be in a correctly protonated (similar to physiological conditions) state. The proton equilibration is done with a suitable hydrogen addition webserver. The prediction of such protonation pKa equilibrium constants of macromolecules can be calculated with the help of atomic resolution of PDB database servers. This can be done with respect to continuum electrostatic models and Lennard-Jones potentials of specific water model force fields. In GROMACS, this can be done with acpype script force field validation and ligand validated with AMBER99SB force fields. The pKa equilibrium constants can be calculated for suitable pH environment with the help of H++ server (<http://newbiophysics.cs.vt.edu/H++/>). This is a key step to build respective calculated pKa dissociation constants for macromolecules (Madeira et al., 2022). In H++ server, these pKa values are quickly estimated to biophysical properties of protein or DNA isoelectric points of respective titration curves at desired pH scales. The required protons are added to protein at suitable sites of various titratable amino acid side chains in PQR format (PDB + Charges + radii). In addition, these pH titration curves and protonation microstates of macromolecules, the topology and molecular model supported AMBER inputs are also generated to create final input PDB for tLeap solvated MD simulated model topology (.prmtop) and input (.inpcrd) files. If the macromolecules are not rendered enough to satisfy the AMBER molecular model package atomic type criteria, the input parameters cannot be generated or while verifying the prmtop files with “cpptraj” and “Parmed” charge validation ended up with error messages. If the MD simulations proceeds with these errors in charged AMBER input parameters, the MD simulations is likely crash at equilibrium and production run phase. Therefore these electrostatic interactions play a key role in ligand binding

to key active site and catalytic and regulatory active site, protein and specified group transfer/ binding to particular regions in macromolecules. These properties of macromolecules vary based upon their specified amino acid side chains protonation states of respective titratable groups, pH and ionic strength of surrounding macromolecules simulated with water sphere.

1.3.8.3 Protein preparation for GROMACS and AMBER MD simulation

Based upon protein-ligand or inhibitor docking scoring functions and docking binding energies and post-docking prediction analysis, the crude approximate ligand/inhibitor three dimensional coordinates are considered to design initial parameters of inputs for MD simulations. But the protein input parameters are not considered like ligand because the protein has different types of residues/groups and their isoelectric point values are assigned with H⁺⁺server (Madeira et al., 2022) and are considered in a pH dependent solvation with periodic boundary limits. The MD simulations of particular protein-ligand system specified with periodic boundary condition as simulation box size is assigned as 8-10 Å. The topology and protein-ligand itp files assigned are system readable with assigned force fields (AMBERff99/14SB) (Hornak et al., 2006) including particular solvation force fields (TIP3P) (Meagher et al., 2003, Mark and Nilsson, 2001). The AMBERff99SB force fields are assigned to ligand with the help of ACPYPE script (Sousa da Silva, & Vranken, 2012) in GROMACS simulations. But in the case of AMBER18 simulations package the atomic charge and positions, angles, torsions are assigned by Antechamber algorithms by employing am1bcc method in AMBER Tools. The entire charge of MD simulations system is neutralized by the addition of Na⁺ and Cl⁻ ions with appropriate charges to the simulation box. In the entire molecular system, the total salt concentration is assigned to be 100 mM. The energy minimization of entire simulated system is carried out for 20,000 steps before entering into temperature and pressure dependent equilibrium of MD system. All positions are restrained at constant temperature (298 °K), pressure (1 atm), achieved with 8 ns MD simulation time before submitting for required production MD run time in AMBER.

1.3.8.4. Minimization and NPT, NVT, Production runs of longer range MD

Energy minimization is a vital step in MD simulations. Because the entire molecular system has various sketched chemical structures, parameters which are not well optimized due to improper bond and dihedral strains are disordered and in random manner. This perturbs the entire system and molecular potential values during MD simulations. As the random molecular system can be aligned chemical motifs stretching and bending, torsional potential energy terms attain the nearest

local minima in minimization step. Therefore, it is a key step to be achieved for local minima to enter next step in equilibration of MD simulations. However, this is not the overall stable total energy of the molecular system as the several algorithms are available to achieve the global minima to attain stability of the system. It is the closest approximate stable minima of molecular structures of entire system with respect to the degrees of freedom.

1.3.9. Basic trajectory data analysis

The post MD simulations analysis can be done using a vast number of module tools some which are inbuilt packages and others are independent of python and C++ based trajectory data analysis. Some of the known python based modules used to investigate basic and advanced MD trajectory data are MDTraj, cpptraj, Pytraj, MDAAnalysis toolkit, g_rms and g_rmsf etc. cpptraj, g_rms and g_rmsf commands based upon GROMACS and AMBER inbuilt trajectory data analysis tool to extract the RMSD and root mean square fluctuation (RMSF) values from MD simulations data. The cpptraj has a larger application and it can read AMBER, GROMACS, NAMD trajectory file formats. The MDTraj and Pytraj, MDAAnalysis toolkit requires a python based graphic library to read the AMBER MD data file formats. Pytraj is a key Python based module useful in dimensionality reduction and extrapolated principal component analysis (PCA) datasets into scatter plots (Hornak et al., 2006).

1.3.9.1. MDTraj and cpptraj, pytraj analysis

MD simulations is a key computational tool for simulations of large biomacromolecules at longer MD simulations timescales. In trajectory analysis, the large data needs special python based module to enhance the advanced trajectory analysis. The Markov state model (MSM) design is a recurrent neural network method needed for higher MD simulations data to monitor the kinetic transition states during micro to millisecond timescales. MDTraj python module provides wide range of MD trajectory data readable space and minimal error in the RMSF and RMSD calculations. However the large number of MD trajectory datasets readable module is essential to design advanced trajectory MD packages with the help of python and C++ programming languages (PyEMMA, Prody). MDTraj is a python based scientific data analysis tool to read large number of trajectory data units including smooth process with wider scientific python environments. It is a bridge between the theoretical molecular simulations data with limit analysed tools to higher order data sets collection to corporate statistical big data analytics including different data visualization machine learning models. It can enhance the shared GPU computation to validate the deep machine learning models with high performance computing cluster to generate a large

dataset. Therefore, the large datasets analysis provides acceleration to understand the biomolecular system and enhance scientific insights in biological process in detail from microsecond bio simulations data.

1.3.9.2. RMSD

It is a relative average positional distance deviation of C α atoms in a protein secondary and tertiary structures while comparing the initial and final backbone movement in MD simulations of structures or comparing an apo protein and protein-ligand complexes. It is a regular parameter measured for protein backbone deviation by superposing atomic coordinates of protein structures from references.

$$\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i=1}^n \delta_i^2}$$

δ_i is the average distance between two successive C α atoms atomic coordinates from superposing ‘n’ pairs of equivalent initial and final protein structures from a given MD trajectory data.

The major flaw in RMSD data analysis is the amplitudes of errors occurring in pairs of identical superposed structures. The lower square distance error deviation is considered as the structural stability in given period of MD timescales. These values range between 2-4 Å of protein in the presence of solvent environment (Kufareva & Avagyan, 2012).

1.3.9.3. RMSF

It is the overall average deviation of protein backbone consisting of amino acid side chains with respect to initial MD structure residue positions.

$$\text{RMSF} = \sqrt{\langle (x_i - \langle x_i \rangle)^2 \rangle}$$

x_i is the position coordinates of initial protein residues.

$\langle x_i \rangle$ is the average positional ensemble coordinates from given trajectory of protein conformers.

RMSF is calculated as the average protein conformational change through the MD simulations and it can identify which region of a protein undergoes conformational deformations in a given average ensemble structures. RMSF is a key parameter to identify the various residues movement in given protein conformers from simulated data sets (Sharma et al., 2021).

1.3.9.4. Binding free energy estimations

Free energy is the nature of energy associated with reaction where the direction of any reaction proceeds that tends to attain equilibrium is thermodynamically favourable. Most of macromolecules like RNA, DNA and proteins undergo conformational changes during their physiological processes. However, the microlevel macromolecular conformational changes occurred in the presence of small chemical molecules (ATP, GTP, FAD, NADPH) to drive the biochemical pathways ensure the cell survival, cell homeostasis, and such processes are associated with free energy changes. Therefore, free energy is an important factor in most of the biochemical reactions to provide the necessary energy to thrive normal physiological process. In general, most of the biophysical reactions are reversible which tend to reverse to initial states to perform the biocyclic process. There are numerous methods to estimate the binding free energies for biomacromolecules. In virtual screening of hit molecules from small molecular databases, to identify the suitable inhibitors according to the protein cavity and scoring functions. These scoring and docking binding energies are calculated as the summation of total atomic and molecular bonding and non-bonding interactions of atomic free energy terms from large macromolecules. Therefore, total free energy is essential to predict the small molecular validation to provide stable protein-ligand complex that leads to the design of viable drug candidates in the near future. In computational drug discovery, small molecular inhibitors are discovered with the help of hit molecular scoring functions (Parenti, and Rastelli., 2012). They are represented in terms of molecular energy surface interactions with protein or DNA cavity energy surface to bind to the active site in order to arrest the macromolecular activity in biological process. The binding free energy can be calculated with continuum solvent simulations for protein and ligand complex designated at periodic boundary conditions (Kollman et al., 2000). The molecular mechanics Poisson Boltzmann or Generalised Born and Surface Area (MM-PB/GBSA) is a widely used method to estimate the reliable free energy path from state A to state B in a given MD simulations (Hou et al., 2011; Srinivasan et al., 1998). It is quite efficient to recognize the molecular paths in host-guest interactions and protein-ligand interactions, protein-protein molecular interactions. However, the quantum chemical method is a more accurate computational method for small molecular structure optimization and simulations in detail (Wang et al., 2017). But most of the biological processes are too complicated and involve thousands of atoms. Therefore the classical simulations are widely used to study conformational dynamics in given explicit solvent models and the molecular free energy is also calculated with implicit continuum solvent models.

Thermodynamic integration and free energy perturbation methods are also explicitly used to estimate overall free energy with umbrella sampling approach before implementing MM-PBSA method (Zwanzig et al., 1954 & 1955; Wang et al., 2017). MM-PBSA approach is robust and reliable to find desirable free energy states in protein-ligand interactions. But it is computationally too expensive in terms of multiple trajectory approach with multiple protein-ligand systems (Wang et al., 2017; Wright et al., 2014; Sitkoff et al., 1994; Gilson, & Honig, 1988).

$$\Delta G_{A \longrightarrow B, \text{solvated}} = \Delta G_{B, \text{solvated}} - \Delta G_{A, \text{solvated}}$$

$$\Delta G_{\text{binding, solvated}} = \Delta G_{\text{complex, solvated}} - \left[\Delta G_{\text{receptor, solvated}} - \Delta G_{\text{ligand, solvated}} \right]$$

$$\Delta G_{\text{solvated}} \cong \langle E_{\text{gas}} \rangle + \langle \Delta G_{\text{solvated}} \rangle - T \langle S_{\text{solute}} \rangle$$

$$= \frac{1}{N} \sum_{i=1}^N E_{i, \text{gas}} + \frac{1}{N} \sum_{i=1}^N \Delta G_{i, \text{solvation}} - \frac{1}{N} \sum_{i=1}^N S_{i, \text{solute}}$$

The above equation represented the MM-GBSA calculated energy terms given from state A to state B. The overall Gibbs free energies differentiate between the selective protein-ligand systems in given states. It has free energy difference among solvated complexes and combined free energy terms of protein and ligands. The MM-PBSA approach has overall solvation energy including the polar non-bonding and dispersion forces interactions added with entropy of protein-ligand complex (Homeyer, & Gohlke, 2012; Wright et al., 2014). MM-PBSA is quite different from MM-GBSA due to the entropy terms calculated including the normal mode analysis (NMA) of overall conformers from MD trajectories in the absence of solvent models. MM-PBSA.py is a python based macromolecule end state free energy estimation approach in the AMBER18 module for molecular simulations and Monte Carlo simulations. Various implicit solvent models estimate the free energy terms for protein-ligand complexes for long range MD simulations (Bennett, 1976; Miller et al., 2012). In Poisson-Boltzmann approach, the vibrational frequencies are calculated to estimate the ensemble based overall entropy of system.

1.3.9.5. Solvent interaction energies

The free energy calculated as the summation of van der Waals and coulombic energies included with reaction fields of cavity effect in solvation free energy. The following equation describes the important energy terms to calculate the interaction energy between host-guest interactions in the absence of solvation energy of molecular system. The solvent interaction energies (SIE) calculated the binding free energy between protein - protein and protein-ligand/inhibitor complexes of MD simulations trajectory data. The interactions considered are similar to MM-PBSA in free energy calculations (Cui et al., 2008). But here the entropy and normal mode energies are not included rather some energy parameters are considered to validate the protein-ligand interactions.

$$\Delta G = \alpha * (E_{\text{vdw}} + E_{\text{Coulombic}} + \text{RF} + \text{Cavity}) + \text{Constant}$$

$$\text{Cavity} = \gamma * \Delta_{\text{SA}}$$

$$\alpha = 0.104758$$

$$\gamma = 0.012894$$

$$\text{Constant} = -2.89$$

This method is quite efficient for protein-protein interactions in both homo dimeric and hetero dimeric solvation free energy of wild type comparative analysis with mutated forms. It calculates the binding affinities similar first order MM-PB/GBSA energy principle but the molecular potentials used are molecular mechanics AMBER force fields continuum solvation models to estimate the binding affinities between protein-protein non-bonding interactions in dimers. However, the SIE is derived entity of MM-PB/GBSA to estimate end point state of the protein-ligand binding energies (Lill, & Thompson, 2011). SIE is mostly depending upon AMBER force field molecular mechanics calculated trajectory data. In this method the number of MD trajectory snapshots used to calculate molecular mechanics average energies sum over to estimate the final binding affinities of protein ligand or protein – protein binding mode (Collins, & Ho, 2019). The SIE method datasets are well matched with most of experimental datasets as well correlated with

other theoretical methods. Therefore, Sie-traj is one of the robust methods to investigate the protein-protein and protein-inhibitor interactions with implicit solvent models.

1.3.10. Python library based advanced trajectory data analysis

1.3.10.1. Prody analysis

Prody is a python based quantitative predictive featured module of structural biomolecules of post MD simulations data. It has combination of different protein quantitative prediction python modules such as principal component square fluctuations, mechanical stiffness, anisotropic network models (ANMs), Gaussian network models (GNMs) and etc (Bakan, Meireles, & Bahar, 2011). The ProDy is a graphical interactive session and a friendly tutorial webserver to predict quantitative measure of protein dynamic motions which can be comparative analysis with experimental results. ProDy is a combination of various other numerous python libraries such as NumPy, SciPy and matplotlib, Scikit-learn which are utilized for deep machine learning models to emphasize dynamic motions from protein –ligand complexes of MD simulations trajectory data. The square fluctuations of individual principal component of anisotropic protein reformative constraints are observed while different types of ligands or substrates were bound at protein active sites. It is one of the machine learning modules to investigate the mechanical deformation induced in proteins by various inhibitors or small molecules that interact with the receptor active sites. It has a combination of different protein sequence trajectory analysis as homologous sequences of wild and mutant forms of various genera of different phylogenetic organisms. The cross correlation plots can be generated among different types of results and comparative theoretical and experimental results of genetic sequence and protein superposition analysis can be done to cross correlate these results in terms of independent PCA. ProDy is run on the MD trajectory datasets of C α atoms, any specific protein domain undergoing refolding and unfolding state datasets which can be efficiently predicted with ProDy python modules. The ANMs can be observed with the help of PCA (Meireles, Gur, Bakan, & Bahar, 2011) and mechanical stiffness from protein-ligand complexes from long range MD simulations timescales. The protein domain motions can be observed and graphical representations can be visualized with porcupine plots extrapolated into two dimensional graphical map. These porcupine plots are generated with the help of uniaxial strength force constants on protein macromolecules binding with various ligands or substrate molecules (Bakan et al., 2014).

1.3.10.2. Principal component analysis

The PCA is a statistical technique to extract the dominant pattern of various structurally distributed representations from unsupervised datasets. The motive behind the higher order multidimensional complex data representation in lower dimension is to investigate the special features with accuracy of individual principal components (PCs) to reduced space in anisotropic changes of system. In computational structural biology, the biomimicry of complex biological systems from real cellular model is carried out with the help of statistical and quantum mechanics ensembles. These simulations provide enormous datasets to conceal the overall information of molecular systems. This might be best correlated data information obtained from PCA to interpret the molecular motions of protein – ligand and protein-protein key interactions from MD trajectory analysis (Eyal& Bahar, 2008). It is applicable to standard reduction of various number of dimensions to exploitable data in two dimensional scatter data plots. Where the dimensionality reduction is a useful tool to monitor the highly dynamic protein motions in a decomposable manner into tiny spatial changes occurring as PCs of respective large datasets. Therefore, higher dimensional datasets are further reduced into a small portions of protein dynamic states to interpret as structural changes in a systematic manner with the help of PCA technique. These small portions of PCA modes reveal the protein dynamic states information at transition state level from hidden kinetic dynamic states from long range timescales. The MD trajectory Cartesian coordinates of protein ensemble data is evaluated into covariance matrix construction with the help of respective eigenvalues deciphered from secular equations. The complete set of orthogonal datasets solved from covariance decomposition eigenvectors show the modes as individual constraint motions of protein anisotropic deformed changes. This can be extracted as the diagonalized of covariance matrix into orthogonal datasets as the secular equations are further validated with respective eigenvalues and eigenvectors to represent the best PCA of individual systems (Yang, Eyal, Bahar, & Kitao, 2009). These datasets are further extrapolated into scatter plots as independent component analysis of respective protein conformations. These square matrix forms of PCA represented with respect to time are called as independent component analysis of specified protein unfolding and refolding events in larger MD timescales.

1.3.10.3. PyEMMA analysis

PyEMMA is an open-source python-based package to construct the high dimensional efficient kinetic state models for long range MD simulations (Scherer et al., 2015). It is reliable to read all the currently available MD simulations formats to enhance the model input selections of design

and investigate model features and validation of stationary states, the dimensionality reduction algorithms such as PCA and time lagged independent component analysis (TICA) (Perez-Hernandez et al., 2013), clustering of MSM estimation and validation with keras algorithms (k-means) (Li, Dong, 2016). The MSM builder is employed to design MSM state model estimation and featurisation, validation done with probable statistical state analogy to further validate the hidden state markov kinetic state model systems to study metastable dynamic state systems. This method provides the pool of statistical validated model systems with Bayesian error calculations for Markov state estimation. It offers various trajectory analysed modules such that the end user can be efficient about molecular trajectory data analysis including molecular observable computations of interest. It can load the large datasets of MD simulation trajectories (up to 100 – 500 GB MD data for 2 nodes (each node 40 processors)) depending upon RAM (500 GB) availability of computing node cluster. The numerous plotting functions are embedded with machine learning model to generate high quality images with the help of Scikit-learn, Pandas, Matplotlib, NumPy, SciPy modules included in PyEMMA (Pedregosa et al., 2011).

1.3.10.4. MSM builder

The MSM builder is a python based package to build the statistical models for long range higher order timescale MD simulations trajectory data. It has larger applications that include time series based biological process and continuous monitoring of the metabolic intermediate in experimental and theoretical predictions to understand the process with respective higher dimensional simulation timescales. To design quasi-dynamic transition states in the complex biological processes, MSM builder (Harrigan et al., 2017) technique can observe the metastable states in most of the biological process is in milli and microsecond time intervals from high order laser and protein NMR spectroscopy. Therefore, the higher order datasets are built from long range biological MD simulations data of protein folding and refolding occurring during the events of enzyme activation and deactivation process and vice-versa. The MSM builder reads the datasets as one or multiple stochastic MD trajectory $C\alpha$ data as atomic positions, distances, angles, torsional and improper torsions angles. It provides a sustainable, reliable and powerful tool to use from scikit-learn API python based results including the high graphical images and dynamic state ensemble PDB datasets from higher dimensions reduction to visualize as scatter plots with machine learning models of existing transition states in a protein-ligand biological simulations. In the current digital age era, tens and thousands GB raw data is generated after MD simulations where these data is analysed by basic trajectory data analytics, the rest of the data is unused.

Therefore, the protein internal dynamics can be extracted with the help of advanced trajectory data analytics such as MSM dynamic state model estimation and dimensionality reductions (PCA, TICA) and further molecular features are enhanced with Perron clustering (PCCA+) validation and transition state path theory (Prinz et al., 2011; Vanden-Eijnden, 2010; Metzner, Schutte & Vanden-Eijnden, 2009). Based upon raw MD trajectory data, the kinetic states designed with the help of dimensionality reduction and random forest, keras iterative mean estimation (k-means) algorithms are used to further build the estimated transition state probabilities to show the further hidden Markov state models associated with microstates. This could facilitate overall quantitative transition probabilities and kinetic populations rates which are relevant to biological phenomenon of macrodynamic states (Shaw et al., 2010; Perez & Granger, 2007; Pedregosa et al., 2011, Noe & Clementi, 2015).

1.3.10.5. Markov state models

The MSMs are a statistical random probability models analysed for dynamic systems. In the long range MD simulations, the trajectory data consists of highly dynamical nature of protein conformational states involved in protein refolding, unfolding and specified loop dynamics. The traditional trajectory analysis cannot identify these high dynamical conformers. The MSMs can be designed by employing MSM builder and then it can be estimated and validated with Bayesian error estimation in order to maximize the random probable stationary states. Further, these MSMs are iterated with model featurisation done to extract pseudo free energies for specified state of protein model (Husic, Pande, 2018). The hidden conformations of protein kinetic states are described as these states to validate with five state model system based upon transition path theory. The neural network model system is used to further express hidden states among kinetic metastable states. These hidden states are essential for protein dynamical states involved in protein allostery, internal structural changes and protein-ligand interactions. This is included in specified allosteric changes in catalytic and regulatory domain in kinase dynamic state distribution among kinetic transitions (Li, Dong, 2016).

1.3.11. Data visualization modules

1.3.11.1. Graphical and plotting modules

The molecular graphics of the three dimensional structures of proteins, protein-ligand and protein-protein interactions were generated with Chimera, Discovery Studio modules (Yang et al., 2012). These modules have also been integrated with other drug discovery modules like virtual screening

of libraries and structural homology Modeller, residue interaction networks (RIN) (Piovesan, Minervini, & Tosatto, 2016), protein sequence homology search module BLAST and multiple sequence alignment fit tools and open source MD simulations visualization software for trajectory analysis using visual molecular dynamics (VMD) and normal mode visualizer are also added as integral part of the graphical modules for data visualization. The protein conformational states are sketched with the help of many graphic modules employing specified applications and python-based advanced graphical libraries, such as Scientific python, Biopython, numpy, TensorFlow, Scikit-Learn, Pandas, Keras-means, Random-forest (Harris et al., 2020; McKinney et al., 2010; Pedregosa et al., 2011; Virtanen et al., 2020; Perez, & Granger, 2007; Cock et al., 2009; Reback et al., 2020; Likas, & Vlassis, Verbeek, 2003). For these analysis high order GPU based graphical drivers are required to concise the unsupervised data into vectorized high order graphical library datasets in two dimensional and three dimensional plotting space.

The modern digital technology has much sophisticated methods to generate high resolution data plots with the help of matplotlib and Scikit-learn, interactive python libraries with advanced python machine learning modules. The big data points can be processed and train the models with the help of deep learning python modules and the data extrapolation can be done with inbuilt plotting python libraries such as logistic regression, Naïve Bayes, posterior probability and prior probability models, unsupervised data model learning, clustering the grouping of data points to minimize the error, data classification with k-means; hierarchical, mean shift, density based clustering of big data.

1.3.11.2. Reduction of dimensionality with neural networks

All machine learning models have been classified into two categories. The supervised and unsupervised data can be analysed and plotted with advanced machine learning python libraries. Data can be interpreted with supervised data analysis with excel and many other plotting methods. This can be done with advanced python source libraries like numpy, Scikit-learn, Pandas data frame. The supervised machine learning models are further divided into regression and classification. The regression models consists of dependent and independent model variables that are further classified in linear regression, decision trees, random forests (ensemble learning methods), neural networks (Hinton, & Salakhutdinov, 2006) (Andrychowicz et al., 2016).

1.3.11.3. The MD traj C α data into dimensionality reduction

The higher dimensionality data can be converted into lower dimensional features without changing high data features to train the multi-layer neural networks. This can be done with gradient descent method by ultra-fine tuning of the data weights which are further deduced into low dimensional centric layer to rebuild high dimensional data vectors described as “antiencoder”. The dimensionality reduction is further categorized into feature elimination and feature extraction. In general, the common method used in simple analysis is PCA. This analysis has overall variance of finding the data points paths along with coordinates which it followed. Therefore the random data points which reduced dimensional data to decrease discrepancy to further reconstruct to back propagate error iteration in encoder network models which is recovered data from decoder called “decoder”. As the data complexity increased, the multi-layer network complex is difficult to be optimized further, if the data points has non-linear encoders and applied gradient methods. The MSM estimation and validation was done with reversible estimation equilibrium transition probabilities with discrete clustering occurring in the random states of protein conformers (Wu, Mardt, Pasquali, & Noe, 2018). The discrete kinetic state models are validated with keras algorithms to further analyse the hidden Markov kinetic models. The implied relaxation timescales are estimated to validate the HMM in order to ensure that kinetic transition probabilities among the microstates during longer range MD simulations. Therefore, the implied timescale analyses and hidden kinetic states indicated that the kinetic transition state distribution occurred in long range MD simulations timescales to reveal the cell physiological transformation possible with specified conformational changes in proteins.

Chapter -2

Enhanced metastable state models of TAM kinase binding to cabozantinib explains the dynamic nature of receptor tyrosine kinases

Abstract

The RTKs are essential proteins responsible for regulation of cell signaling. The atomistic details and mechanism of functional regulation is required to understand their normal physiological processes and when overexpressed in disease conditions. TAM RTKs are inhibited by cabozantinib at nanomolar concentrations. The docking of cabozantinib, a nanomolar affinity inhibitor in the active state conformations of TAM kinases revealed the best binding pose and the complex formation mediated through non-bonding interactions involving the hinge region residues. The alterations in the conformations and the regions of flexibility in apo and complexed TAM kinases as a course of time are studied for 250 ns using molecular dynamics (MD) simulations. The stabilized trajectories in all molecular systems were analysed using post-MD analyses to reveal the dynamic and active metastable states. The encrypted protein domain motions were analysed using various post-trajectory analysis tools like ProDy and MDTraj. The large scale motions in the cabozantinib bound Tyro3 and Mer is indicated by large eigenvalues. Square fluctuation plots based on principal component analyses revealed more populated microstates and lower displacements in Axl revealing its structural stability. Mechanical stiffness plots revealed that similar regions in apo and cabozantinib bound Axl fluctuated during MD simulations whereas different regions in Tyro3 and Mer kinases, indicative of the differences in the conformational stabilities of proteins. The RIN plots revealed the important salt bridges that lead to constrained or limited domain motions in the TAM kinases.

2.1. Introduction

RTKs are essential trans-membrane components in cellular signal transduction pathways (Lemke, 2013). One subfamily of RTKs consists of three proteins, Tyro3, Axl and Mer collectively called as “TAM RTKs”. The extracellular factors such as Gas-6, Pros1 and EGF, activate TAM kinases (Akalu, Rothlin, & Ghosh, 2016). TAMs play crucial roles in a variety of normal biological functions such as spermatogenesis, bone physiology, controlling platelet aggregation, endothelial and vascular smooth-muscle homeostasis (Gay, Balaji, & Byers, 2017; Vouri, & Hafizi, 2017). TAM RTKs are also reported to play crucial roles in disease conditions such as acute myeloid leukemia, breast, colorectal, lung, ovarian cancers and glioblastoma (Rankin, & Giaccia, 2016; Knubel et al., 2014). The three dimensional (3D) high resolution crystal structures of TAM kinases have been reported with PDB_ids; Tyro3 (3QUP), Axl (5U6B) and Mer (5TC0, 5U6C, 2P0C, 5TD2). With the exception of 5U6B_B and D chains (Gajiwala et al., 2017), all TAM kinase structures solved to date are in the inactive conformation. The active state conformations of Mer and Tyro3 and the conformational alterations in the ensemble of active states in TAM kinases is not revealed. In this work, the active state forms of TAM kinases from crystal structure and homology models were analysed. Binding of high affinity inhibitor, cabozantinib has been studied using molecular docking and MD simulations. Further, the MD trajectory files were analysed in detail with MDTraj, MDAnalysis and ProDy Python libraries to study the dynamic microstate analysis and extended the model building with MSM (Sultan, Kiss, & Pande, 2018). This work reveals the ensemble conformational microstates of active TAM kinases from MD simulations studies.

2.2. Materials and methods

2.2.1. TAM kinase structures and Homology modeling

The multiple sequence alignment of the kinase domains in TAM RTKs was constructed using Clustal Omega online server (Sievers et al., 2011). The active state conformation of Axl (PDB ID: 5U6B_B chain) was considered and the three missing residues in the activation loop were built using "Model/Refine Loops" in "Structure Editing" tool in UCSF Chimera 1.11 (Yang et al., 2012). The active state model structures of Tyro3 and Mer were constructed based on the template structure of 5U6B_B chain. Homology modeling was carried out using MODELLER 9.21 method implemented using Python script (Šali, & Blundell, 1993; Webb, & Sali, 2014). The quality of Mer and Tyro3 model structures was validated using structure validation servers such as PROCHECK (Ramachandran, Ramakrishnan, & Sasisekharan, 1963; Ramachandran, & Venkatachalam, 1968; Laskowski, Macarthur, Moss, & Thornton, 1993), Verify_3D (Bowie, Luthy, & Eisenberg, 1991) and ERRAT (Colovos, & Yeates, 1993).

2.2.2. Molecular docking

The coordinates of cabozantinib were deduced from the Ephrin type-A receptor 2 RTK-foretinib complex (PDB_ID:5IA4) (Heinzlmeir et al., 2016). AutoDock4 (Morris et al., 2009), a molecular docking tool was used to dock cabozantinib into the ATP binding site of TAM kinases. AutoDock is an open source software to study the protein-ligand conformations, best suited to predict the ligand conformations in the binding site environment of a protein. It is a grid-based docking technique developed upon Lamarckian genetic algorithm to generate global poses of small molecules. The empirical free energy force field that includes hydrogen bond direction term with explicit polar hydrogens and contribution from electrostatic interactions is used to quantify the binding free energies of the docking poses. In 5U6B, the ATP binding site is occupied by the co-crystallized small molecule, a macrocyclic inhibitor. The grid parameters are fixed at this binding site that includes hinge region for docking of cabozantinib (box parameters: X: 50, Y: 36, Z: 40), spacing (0.408 Å); center of grid (X, Y, Z) was defined as (29.097, 3.263, 52.193). The conformation with best binding energy and maximum docking poses was considered as the best binding conformation and was proceeded to further MD simulations.

2.2.3. Molecular dynamic simulations

The MD simulations of apo-TAM kinases and when bound to cabozantinib were studied using GROMACS 5.1.4 version (Van Der Spoel, 2005) for 250 ns. The MD simulations force field parameters for proteins were derived by using AMBERff99SB (Hornak et al., 2006). The cabozantinib force fields were generated in Antechamber using AM1-BCC method in ACPYPE script (Sousa da Silva, & Vranken, 2012). All molecular systems were placed in a 3D cubic box and were solvated using SPC waters as single point charge (Mark, & Nilsson, 2001). In order to neutralize the systems, eight Na⁺ ions were added throughout MD simulations. Energy minimization was carried out by using steepest descent method for 50,000 steps to overcome short range bad contacts (Fletcher, & Powell, 1963). Further, these molecular systems were proceeded for equilibration and production phases. In the equilibration phase, position restrained simulations were performed for 1 ns at 300 K under NVT (constant number of particles, volume and temperature) called the "canonical ensemble" at constant temperature. Subsequently, in the next equilibration phase, the pressure of the system was stabilized at 1.41 bar under NPT ensemble (constant number of particles, pressure and temperature) called the "isothermal-isobaric" ensemble (Hess, Kutzner, van der Spoel, & Lindahl, 2008). In the production phase, the final MD simulations were carried out for 250 ns at 298 K temperature and 1 bar pressure, using a time step 0.002 ps. The V-rescale thermostat couple was used to

monitor the temperature and Parrinello-Rahman method was utilized to maintain pressure (Bussi, Donadio, & Parrinello, 2007). Particle Mesh Ewald method was used to maintain long range electrostatic interactions with real space cut-off of 10 Å and it has an order of 4 (Darden, York, & Pedersen, 1993). The relative tolerance cut-off range energies were 10^{-6} among short, medium and long-range interactions. Short range interactions were calculated including neighbor list of 10 Å to update after every 10 steps. The van der Waals forces were estimated using Lennard-Jones potentials and real space electrostatic interactions was truncated at 9 Å (Verlet, 1963). Linear Constraint Solver (LINCS) algorithm was utilized for scaling hydrogen bonding constraints (Hess, Bekker, Berendsen, & Fraaije, 1997). After MD simulations, the snapshots of average structure were extracted from GROMACS trajectory file (.trr) and were visualized using VMD, a molecular visualization program. The conformational analysis of TAM kinase – cabozantinib complexes were analyzed using `g_rms` and `g_rmsf` commands for RMSD and RMSF plots, respectively. The average structures were obtained from the 250 ns of MD simulations data. The convergence of MD production run was explained in terms of potential energy and RMSD plots of the molecular systems. Dynamics of the apo and protein-inhibitor complex structures was taken from 500 snapshots of 250 ns MD simulations data.

2.2.4. Post-MD data analysis

Trajectory data points were derived from .trr files or .xtc files of the 250 ns MD simulations run using GROMACS 5.1.4. MDTraj is a software package that allows users to manipulate MD trajectories from a variety of file formats to simplify the analysis of MD data and connects the datasets with the software in Python (McGibbon et al., 2015). The protein structural dynamics are characterized from domain motions and structural features could be discretized using various python programming libraries. The data sets derived from trajectory analysis are used for understanding the key conformational states in TAM kinases. The domain motions in protein can be monitored with the help of PCA (Meireles, Gur, Bakan, & Bahar, 2011). These can be achieved with the application of programming interface using Python based “ProDy” (Bakan, Meireles, & Bahar, 2011). Python applications are a combination of various plugins like SciPy, NumPy and Matplotlib which help to plot data of derived covariance and eigenvalue, eigenvectors. The datasets are plotted as principal components of structural domains of protein in terms of calculated covariance matrices of small constraints for principal modes in various parts of the protein such as α -helix, β -sheet and loop variations. The small constraint anisotropic motions are observed due to variations in principal protein conformations in MD simulations studies. The individual constraints of modes show normal mode analysis (NMA) in graphical interfaces such as anisotropic normal modes (ANM) (Bakan et al., 2014).

ANM can be derived as individual modes which include diagonalized Hessian matrix, scrutinizing eigenvectors and calculated subspace overlap of constraint motions in protein. The mechanical stiffness measures the mechanical resistance to external force that is applied at specific pairs of residues on the 3D structure or a complete map of the mechanical resistance in response to all possible pulling directions. The ANM of the simulated structure is used as an input to generate a complete 2D map of mechanical stiffness/strength as a function of residue indices. The effective stiffness or force constant for each residue averaged over all pairs of residues formed is plotted as 1D profile. The mechanical stiffness is an indirect method to examine the uniaxial strength in protein-inhibitor complexes. In the long-range MD simulations, the protein undergoes partial folding and unfolding in the absence/presence of inhibitors. The protein unfolding pathway proceeds via the uniaxial forces applied along anisotropic directions which make specified mechanical or elastic modes in protein-drug complex. These are numerical insights to predict the mechanical resistance of protein models in a specified residue range under the influence of inhibitors fitted into the protein active site (Eyal, & Bahar, 2008; Chen, Wang, & Zhu, 2016). The active state distribution and structure featurization, clusterization, TICA and MSMs are extracted and discretized using pyEEMA 2.5.5 python package (Lane, Bowman, Beauchamp, Voelz, & Pande, 2011; Scherer et al., 2015; Li, & Dong, 2016; Husic, & Pande, 2018). These MSMs are kinetics models and measure the metastable forms which are related to molecular kinetics. This kinetic model estimation is done with selective input features of simulation dynamics using a dimensionality reduction algorithm, tICA. Specified kinetic model sketches have been done by using clustering methods such as k-means and further derivatized into MSM models.

Principal components in variable domain motions of proteins can be visualized and extrapolated as porcupine plots of individual protein dynamic motions. The elucidation of 3D array of non-bonding interactions and different types of bonding contacts among amino acids in protein structure is vital to define its function. The RIN analyzer plays a vital role to describe and visualize these interactions. The RING 2.0 web server gives these preferable interaction networks among main chain atoms in a protein structure (Piovesan, Minervini, & Tosatto, 2016). The 3D structure of a protein can be visualized in a 2D plot by mimicking of high dimensionality and virtualizing of plot similar to 3D. These plots are generated using "Cytoscape" as interface connected with chimera through "structureViz" (Morris, Huang, Babbitt, & Ferrin, 2007). RIN network parameters can be customized with RIN analyser app wizard, amino acids are represented as "Nodes" and interactions as "Edges" in Java application-based interface Cytoscape (Doncheva, Klein, Domingues, & Albrecht, 2011).

2.3. Results and Discussion

2.3.1. Sequence analysis

Tyros3, Axl and Mer are homologous RTKs. The multiple sequence alignment of the kinase domains in the TAM kinases is shown in Figure 2.1A. These proteins share high structural conservation in the kinase domain with greater than 62% sequence identity. The TAM kinase crystal structures available in PDB are in the inactive state Mer (5TC0, 5U6C, 2P0C, 5TD2) and Tyro3 (3QUP), with the exception of Axl. The crystal structure of Axl kinase domain (5U6B) is in two distinct conformations; subunits A and C are in the inactive state, while subunits B and D are in the active state (Gajiwala et al., 2017). The active state structure of only Axl kinase among TAMs is available; we initially added the missing residues (705, 706 and 707) in the activation loop of Axl_B chain and used this as a template for constructing the homology model structures of Mer and Tyro3. Among the five distinct models of Mer and Tyro3, the model structures with lowest DOPE scores were selected (Shen, & Sali, 2006). The statistical parameters for the model validation are shown in Table 2.1. These parameters indicated that the constructed models of Tyro3 and Mer are suitable for further studies. The structures of TAM kinases in their active state superposed well as shown in Figure 2.1B.

Phosphorylation and dephosphorylation of the kinase domain allows the conformational switch from active to inactive forms and vice-versa (Levinson et al., 2006). From the available crystal structures and analyses of c-Src and Abl (Meng, & Roux, 2014), the essential structural differences between the active and inactive states of a kinase are reported. The main structural differences are the altered conformations of α C-helix in the N-terminal domain, orientation of the DFG motif and opening of the activation loop. The activation loop in the kinase domain plays a major role in autophosphorylation and ATP binding. In the active state, activation loop is phosphorylated, opened up with an outward extended orientation. A catalytically important conserved DFG motif at the base of the activation loop has Asp side chain pointing inwards into the active site. Inward movement of the α C-helix, with the side chain of conserved Glu pointing inwards and the conserved Lys on the β 3-strand in the N terminal domain make ionic interactions within the active state. Distinct conformations for the inactive states of kinases are similarly revealed, with closed conformation of the activation loop, outward movement of the Glu side chain from α C-helix and Asp side chain from DFG motif (Shukla, Meng, Roux, & Pande, 2014). In the Axl kinase crystal structure, the proposed autophosphorylation sites (Tyr698, 702 and 703) in the activation loop are not phosphorylated in spite of the structure being in the active state conformation. In the Axl kinase active state conformation (5U6B_B), E585 (α C-helix), D690 (DFG motif) and K567 (β 3-strand) are pointing inwards and make ionic

interactions Figure 2.1C. These interactions provide a good validation for the active state conformation in the kinase structures. In the inactive conformation (5U6B_A), these residues move away and the side chains of E585 and D690 flip and orient away from the active site (Figure 2.1C). We have investigated the alteration or retention of these interactions with help of atomistic MD simulations studies in TAM kinases.

2.3.2. Molecular docking of cabozantinib

The primary protein-ligand complex formation can be approximated with molecular docking studies which would find probable cabozantinib binding in the active state of TAM kinases. In the molecular docking studies carried out using AutoDock4, the docking pose of inhibitor in the protein active site is in a horizontal manner and is located in the ATP binding site as shown in Figure 2.1D. The binding energies of the best docked pose are shown in Table 2.1. The binding site is fully engaged with inhibitor, where the activation loop has extended interactions with the active state of TAM kinase. As shown in Figure 2.1D, the binding of cabozantinib to TAM kinases can be divided into three chemical regions. The primary part of the inhibitor, (6, 7 dimethoxy quinoline) containing nitrogen interacts with hinge region main chain of Phe622 and Met623 residues (Axl), (Mer: Phe666 and Met667) and (Tyro3: Phe607 and Met608). The other end of the inhibitor consists of an amide linked p-fluoro phenyl that is π stacked with Phe691 (Mer: Phe735 and Tyro3: Phe676) aromatic ring which is a part of DFG motif, and the amide carbonyl linkage with cyclopropyl ring interacts with Asp690 which is also a part of DFG motif (Mer: Asp734 and Tyro3: Asp675). The amide linked fluoro phenyl ring is immersed into a deep cavity which is in the neighborhood of α C-helix in TAM kinases. A three-residue junction networked via non-bonding interactions from Lys567 (β 3-strand), Glu585 (α C-helix) and Asp690 (A- loop) (Axl numbering) is in the inhibitor binding site and the inhibitor gains access through these residues in all TAM kinases



Figure 2.1A) Sequence comparison of Tyro3, Mer, Axl RTKs.

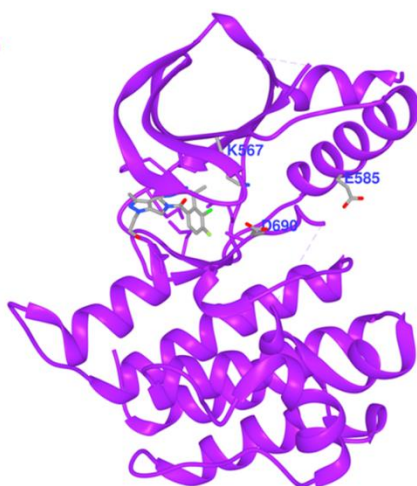
(P loop) Lys – Glu (α -helix)

(P loop) Lys – Asp (DFG Motif)



Figure 2.1B

Axl-Active 5U6B Chain B



Axl-Inactive -5U6B Chain A

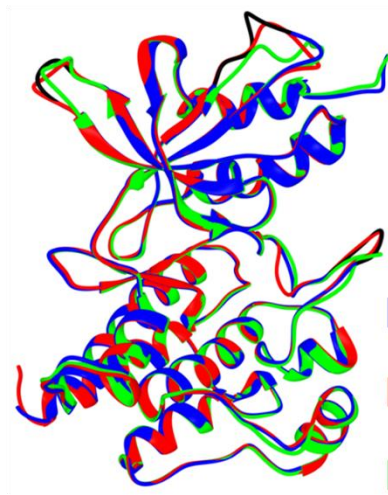


Figure 2.1C

Superposition of TAM kinase domain

■ Axl
■ Mer
■ Tyro3

Figure 2.1B) Structure superposition of kinase domain in TAM RTKs. Green – Tyro3 model; Red – Mer model; Blue –Axl 5U6B; Missing residues added in Axl are shown in Black color.

Figure 2.1C) Active (5U6B_B, magenta) – Inactive (5U6B_A, violet) state conformations in the Axl crystal structure. The three residue interaction network between the side chains of Lys567-E585 and K567-D690 are indicated.

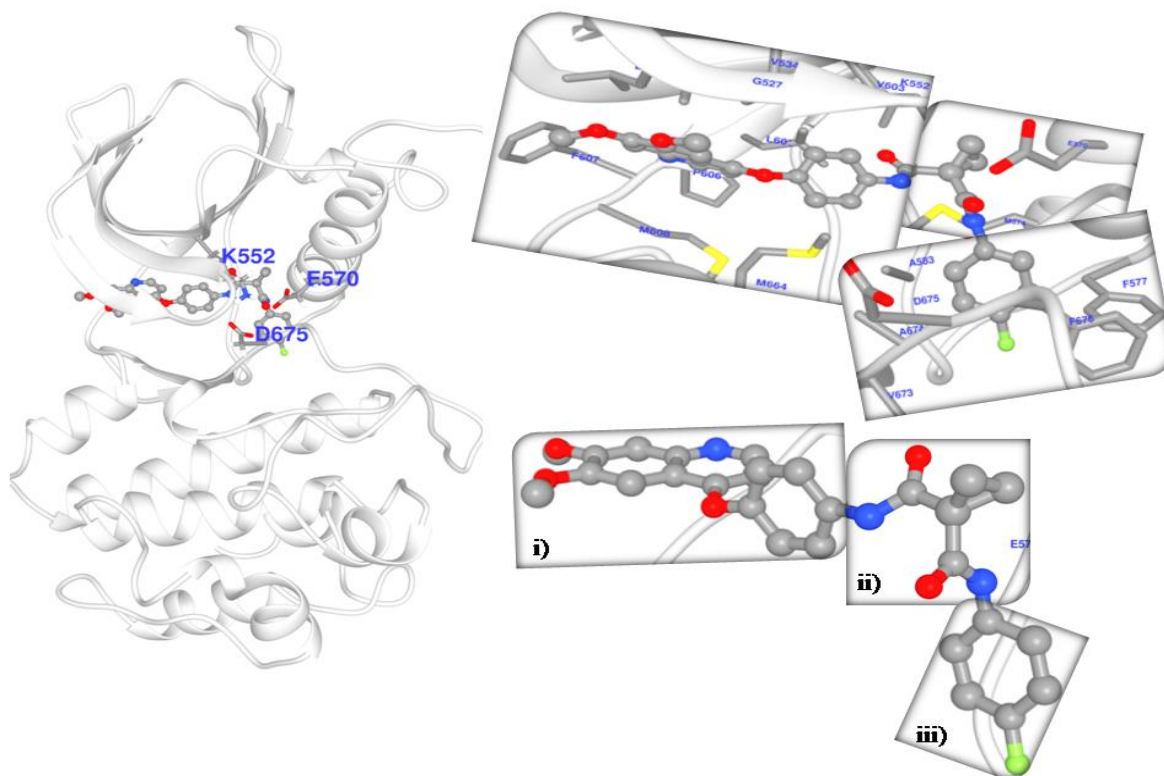


Figure 2.1D) Structural comparison of docked cabozantinib in Tyro3 receptor kinase.

Tyro3 (White); Cabozantinib (elemental color); i) 6,7 dimethoxy quinoline ; ii) cyclopropyl stacked amide carbonyl; iii) fluoro phenyl ring.

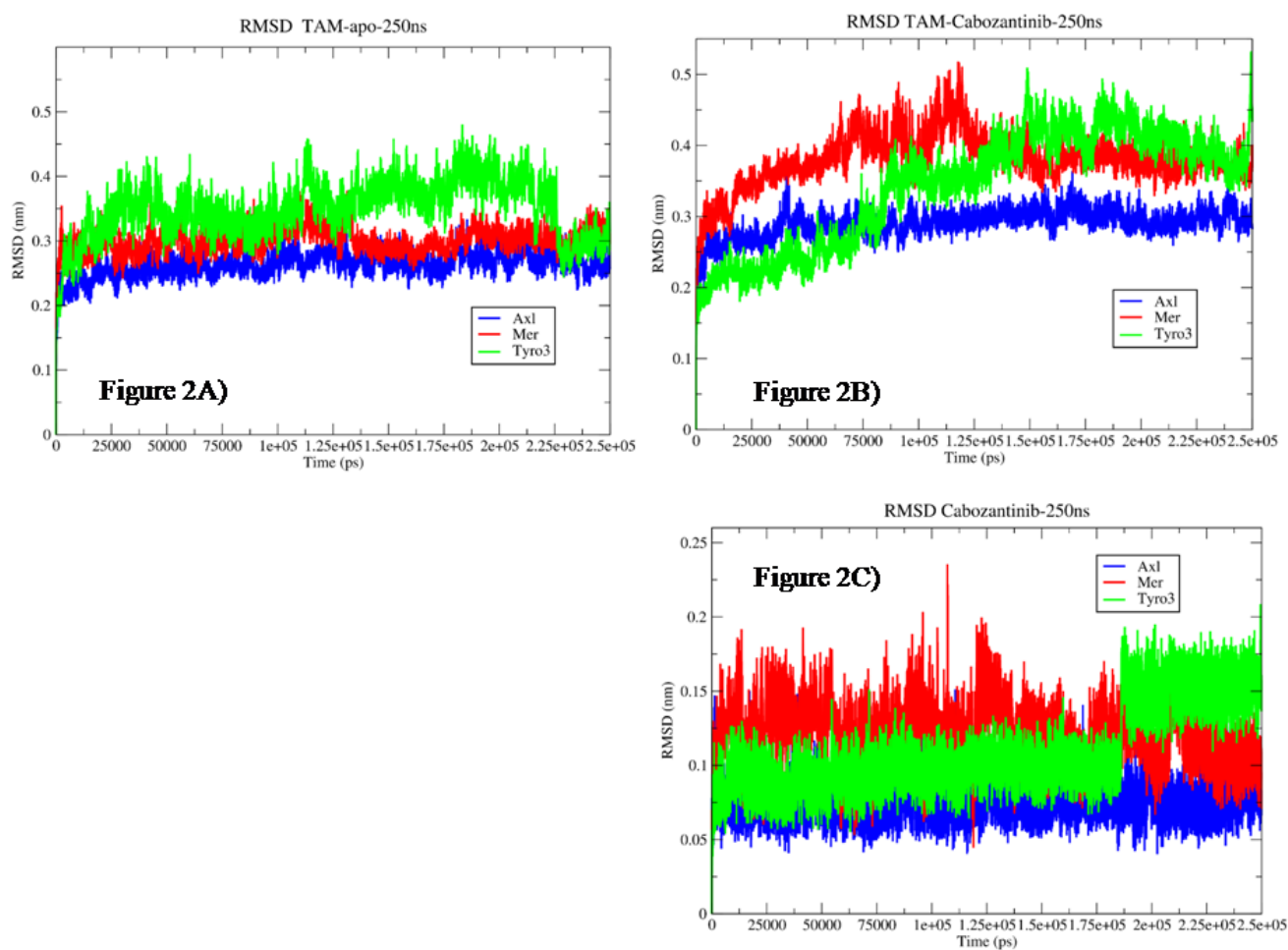


Figure 2.2) RMSD plots of apo and Cabozantinib bound TAM RTKs from 250 ns MD simulations. (A) Apo proteins. (B) Complexed protein (C) Cabozantinib.

2.3.3. MD simulations of apo and TAM kinase- cabozantinib complexes

Six molecular systems; Tyro3, Axl and Mer kinases; apo and complexed with cabozantinib were studied by MD simulations using GROMACS 5.1.4 for 250 ns. All MD simulations converged well as indicated by stable potential energies that are shown in Table 2.1. The RMSDs of apo and TAM-cabozantinib complexes are compared in the given stable proteins relative to the average distance in the protein backbone. It is one way of measuring the change in average C α distance which influences certain parts of the protein structure. The overall RMSD of all molecular systems are stable as shown in Figure 2.2 (A, B, C) these plots indicated that all the molecular systems have converged after 10 ns of MD simulations. In the apo Axl (Figure 2.2A) and when bound to cabozantinib, the RMSD of the C α atoms is relatively lower which is indicative of the greater stability in their structures. On the contrary, apo Mer is more stable than inhibitor bound form. Both apo and inhibitor bound Tyro3 display higher RMSDs in comparison to other molecular systems. This implies that there are more possible metastable

active states existing in the case of Mer and Tyro3 kinases compared to Axl. Therefore, the RMSD of both Mer and Tyro3 kinases are slightly higher in a fortuitous manner to expect their dynamic nature from MD simulations data. This could be a key aspect from protein dynamics that plays a vital role in the drug design of TAM kinases in the identification of allosteric sites for new inhibitor design. When TAM kinases are bound to cabozantinib, they have specific states, which are catalytically active, providing insights for atomistic pinning of TAM kinase activity.

S.No		Mer-Model	Tyro3-Model	Axl (5U6B_B)
1.	Protein model validation Model			
	DOPE Score (Modeller)	35206.7775	34680.4453	-
	ERRAT (Overall quality factor)	78.24 %	76.84 %	-
	Verify_3D (avg 3D-1D score >= 0.2)	80 % aa pass	80 % aa pass	-
	ProSA(Z-Score)	-8.44	-7.79	-
	PROCHECK Ramachandran plot			
	most favoured	92.7 %	91.5%	
	additional	6.2 %	7.3 %	-
	allowed	1.2%	1.2%	
	generously allowed			
2.	Docking binding energy (ADT) in k.cal/mol	-8.93	-7.15	-7.06
3.	Potential Energy of 250ns MD simulated system (k.cal/mol)			
	Apo protein	- 4.256 * 10 ³	- 4.851 * 10 ³	- 4.737 * 10 ³
	Cabozantinib complex	- 4.219 * 10 ³	- 4.853 * 10 ³	- 4.457 * 10 ³

Table 2.1: TAM kinase models, cabozantinib docking and MD simulations validation

	TAM with Inhibitor form	
3 Residue number	50ns Å	250 ns Å
Axl K-567-E-585-D-690	Met-623 (Hinge) Asp-690 ; Phe-691 (DFG)	No H bond
Mer K-612-E-630- D-734	No H bond	Met-667 (Hinge)
Tyro3 K-552-E-570- D-675	Met-608 (Hinge) Glu-570 (α C-Helix) Phe-676 ; Gly-677 (DFG)	Met-608 (Hinge) Lys-552 (P-Loop) Asp-675 (DFG)

Table 2.2) The three residues interactions in TAM kinase bound cabozantinib for 250 ns MD simulations data.

The MD simulations explain the dynamics and alterations in the conformations of TAM kinases at different time scales of apo and inhibitor bound forms as a course of simulations time. During MD simulations, it is interesting to see the appearance of new non-bonding interactions tightly binding the ligand in the protein cavity. The trajectory analysis of 250 ns of MD simulated data was analyzed for scrutinizing the non-bonding interactions among the specified active states of the protein structures. The regions from the N-terminal domain (β 3-strand, α C-helix) and activation loop are involved in strong interactions in all the active structural conformers of TAM kinases. In the complexed TAM kinases, these distances have moved slightly away, however, the side chains of the amino acids are pointing inwards. To understand this in detail, the distances between the atoms involved in forming the three residue interactions in TAM kinases were measured and are shown in Figure 2.3 (A, B). These distance plots indicate the retention or loss of hydrogen bonds in the core of kinase domains. The plots of distances between the two electronegative atoms participating in the hydrogen bonds were quantified by plotting pie diagrams as shown in Figure 2.4 (A, B, C). In the apo Axl MD simulations, Lys567 N ϵ and Glu585 O δ , retained the hydrogen bonding distance between 2.5 to 3.4 Å upto 92% times during the 250 ns MD simulations. Similarly, in the cabozantinib bound form, the distance between these atoms is retained upto 90.4% times. The distance between Lys567 N ϵ and Asp690 O γ is also retained upto 34.9% in the apo Axl and 30.6% in the cabozantinib bound Axl.

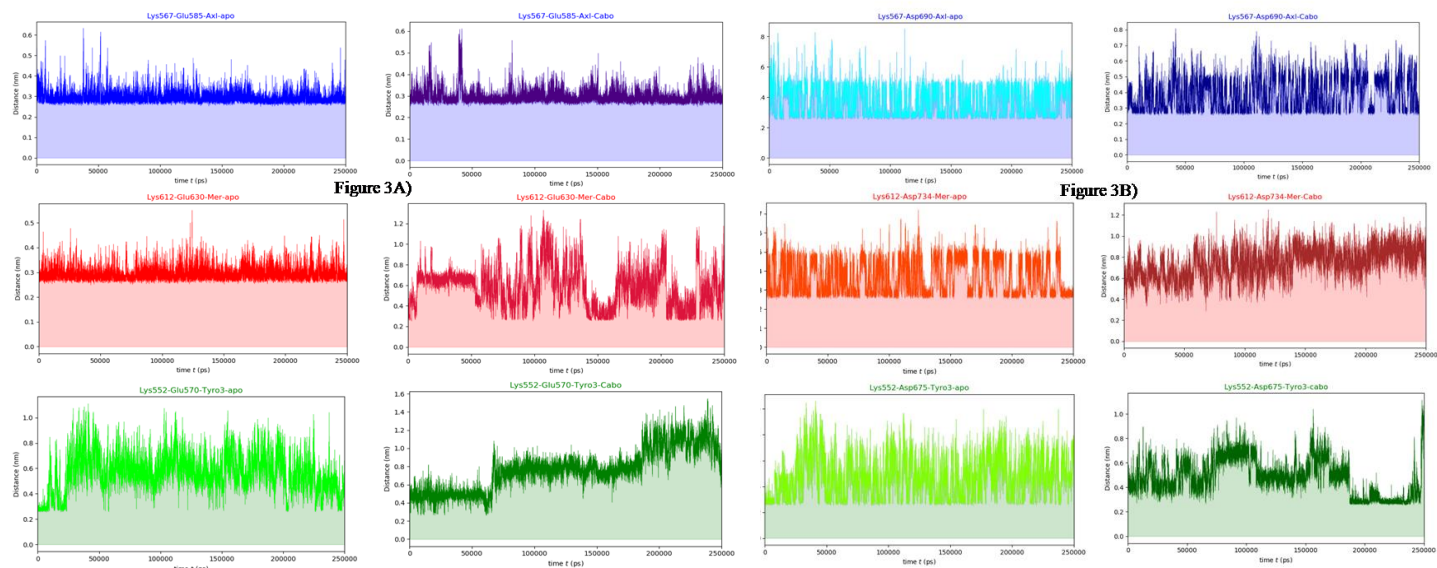


Figure 2.3 (A, B) Average distance between the side chains of Lys-Glu and Lys-Asp in the 250 ns MD simulations in TAM RTKs.

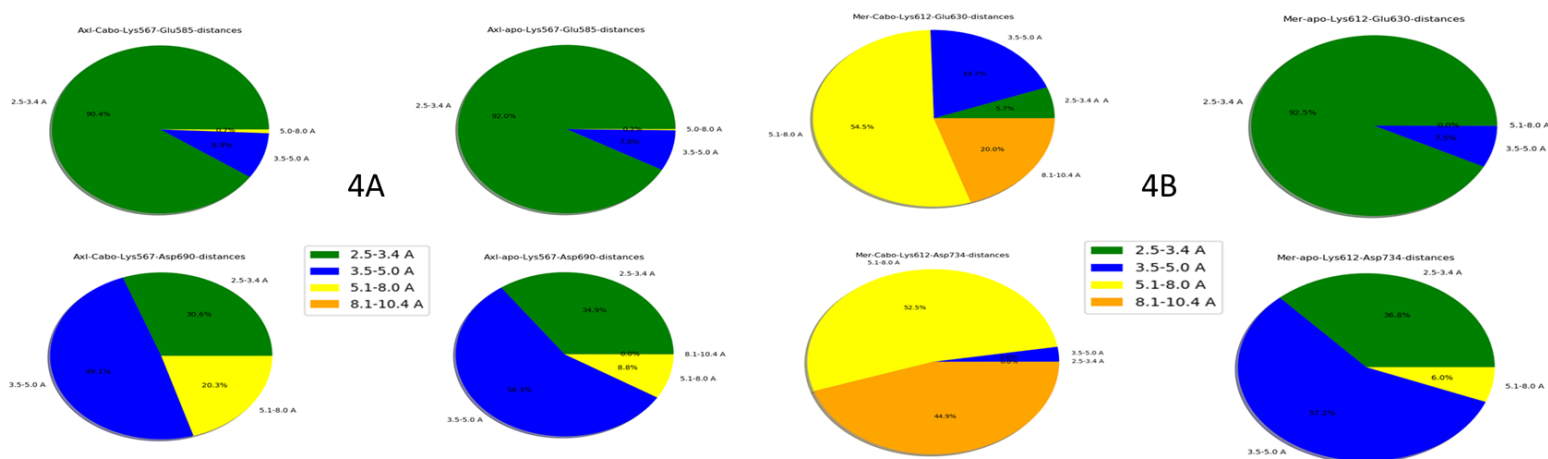


Figure 2.4A) Distance between side chains of Lys - Glu; Lys - Asp pairs in apo and Axi-mercabozantinib RTK from 250 ns MD simulations.

Figure 2.4B) Distance between side chains of Lys - Glu; Lys - Asp pairs in apo and Mer-mercabozantinib RTK from 250 ns MD simulations.

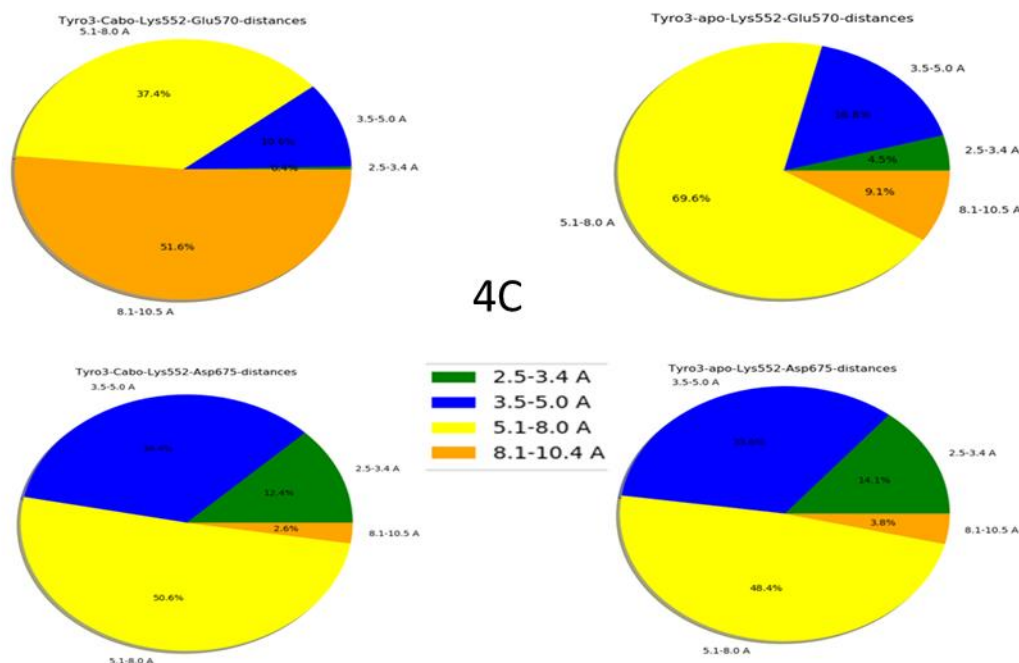


Figure 2.4C) Distance between side chains of Lys - Glu; Lys – Asp pairs in apo and Tyro3-cabozantinib RTK from 250 ns MD simulations.

In the case of apo Mer kinase, the hydrogen bonding distances between Lys612 N ϵ and Glu630 O δ is retained upto 92.5% times, whereas, in the cabozantinib bound form this distance is retained only in 5.2% and the distance drastically increased upto 10.5 Å. The Lys612 N ϵ and Asp734 O γ distance in the apo Mer kinase is also moderately retained upto 36.8% and in the cabozantinib bound Mer, this interaction is completely lost and the distance increased upto 10.4 Å during the 250 ns MD simulations. In the case of apo Tyro3, the hydrogen bonding distance between Lys552 N ϵ and Glu570 O δ is retained only upto 4.5% and became almost insignificant (0.4%) in the cabozantinib bound Tyro3. In both apo and cabozantinib bound Tyro3, the Lys552 N ϵ and Asp675 O γ distance was retained upto 14.1% and 12.4% times, respectively. The pie diagrams of these hydrogen bonding distances indicated that the three residues interactions are retained in the apo Axl and Mer, and cabozantinib bound Axl, indicating that the active site core of the protein is retained only in these molecular systems. The expansion of the active site core of the other three molecular systems (apo Tyro3 and cabozantinib bound Mer and Tyro3) is indicated by the loss of hydrogen bonding interactions.

From the RMSF plots, shown in Figure 2.5, we see that similar regions in all the molecular systems of TAM kinases fluctuate during MD simulations. In the case of apo TAM kinases the extent of fluctuations is almost similar whereas in the cabozantinib bound forms the extent of

fluctuations varies in specific regions. For instance the regions; a β -turn in the N-terminal domain) and end of first α -helix in the C-terminal domain have least fluctuations in Axl complex, indicating that cabozantinib binding to Axl stabilizes this region. Similarly the activation loop has relatively higher fluctuations in cabozantinib bound complexes compared to the apo form. These differences in the RMSF plots indicate the extent of conformational flexibility in TAM kinases and their structural plasticity. The differences observed in the RMSF plots can be explained by analysing the trajectory files of MD simulations. The MD simulations data clearly reveals that there are minute changes occurring at specific regions with reference to longer MD time scales. The changes in the protein conformational features with respect to simulation time scales can be explained from the analyses of the MD trajectory files.

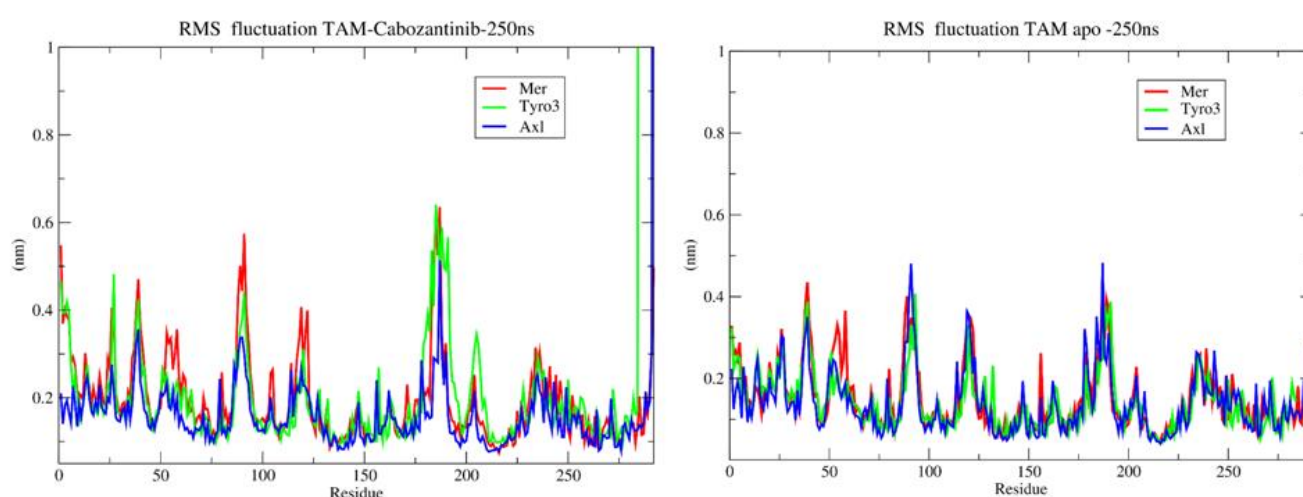


Figure 2.5) RMSF plots of TAM RTKs from 250 ns MD simulations. (A) Cabozantinib complexes (B) Apo proteins. In this figure, for the sake of convenience all TAM kinases are numbered from number one (1) onwards. 88-93 (β -turn in the N-Terminal domain); 117-122 (First α -helix in the C-terminal domain); 179-192 (activation loop).

In the case of cabozantinib bound to TAM kinases at different time scales of MD simulations, the orientation of both Mer and Tyro3 change due to movements in α C-helix and activation loop deformations. Whereas the Axl MD simulations data indicated only minor changes or little movement in α C-helix in the apo and the cabozantinib bound forms. The distances for the non-bonding interactions among (Axl: K567, E585 and D690) is given in Figure 2.4. In Mer and Tyro3 these interactions are quite weak due to the elongation of the corresponding

distances resulting from the α C-helix and activation loop dynamics. These dynamic states are closely associated with each other and can change specified non-bonding residue interactions. The α C-helix of apo Mer kinase is similar to the apo Axl, however the cabozantinib bound Mer α C-helix moved away with local unwinding of helicity. Both apo and cabozantinib bound Tyro3 display changes in the α C-helix and movement away from the core of the protein. Also this figure indicates that the interactions between the N-terminal Lys and activation loop Asp (DFG motif) is retained only 1/3rd times in Axl and are significantly lost during the MD simulations of Mer and Tyro3. It is the major diversity in the α C-helix and activation loop conformations that causes generation of more metastable state populations in Mer and Tyro3. These states can be explained with advanced trajectory data analysis by Python programming libraries. The population analysis of these conformations is further explained with post-MD data analysis from longer timescales MD simulations trajectories.

2.3.4. Post-MD Analysis

The MD simulated TAM-cabozantinib systems generated large datasets which are difficult to be inspected using pure visualization software for structural alterations and protein domain motions. Preliminary post-MD data analysis was carried out on the large trajectories derived from TAM kinase 250 ns MD simulations data using “ProDy” Python libraries to ensure kinetically active states are investigated with the aid of PCA. PCA is a powerful multivariate statistical technique to decipher the conformational changes in a protein as a function of time, this was therefore used to study the MD simulations trajectories of TAM kinases. PCA is applied to systematically reduce the number of dimensions needed to describe the protein dynamics through a decomposition process that filters the observed motions from largest to smallest spatial scales. The presence of large-scale motions makes it difficult to reveal the small-scale motions in proteins since the large-scale motions have relatively greater amplitude in atomic displacements. Therefore, only a small number of PCA modes with large amplitudes that reveal the structural dynamics of biological relevance are considered. A covariance matrix is constructed from the Cartesian coordinates of the trajectory files and an eigenvalue decomposition of the covariance matrix leads to a complete set of orthogonal collective modes (eigenvectors), each with a corresponding eigenvalue that characterizes a portion of the motion. The eigenvalues were obtained by the diagonalization of the covariance matrix of the C_{α} atomic fluctuations and depicted in Figure 2.6, with decreasing order versus the corresponding eigenvector indices. The square matrices are converted into secular equations to generate eigenvalues and corresponding eigenvectors. The corresponding eigenvectors are the best principal components associated with eigenvalues (Yang, Eyal, Bahar, & Kitao, 2009). The

first few eigenvalues corresponding to concerted motions quickly decrease in amplitude to reach a number of constrained and more localized fluctuations. As shown in the Figure 2.6, the eigenvalues of apo Axl and when bound to cabozantinib are low and almost similar. While the Tyro3 and Mer bound to cabozantinib have higher eigenvalues compared to apo Tyro3 and Mer. This suggests that the nature of motions in all the six molecular systems is quite different, however the extents of dissimilarities are lower in apo and cabozantinib bound Axl. Larger eigenvalues typically indicate large scale motions spatially, implying that the conformational alterations in the cabozantinib bound Tyro3 and Mer are greater during the course of MD simulations. The trajectory data sets are converted into a covariance matrix to generate 10 principal components as shown in figures 2.7 (A,B,C). This figure shows square displacements of TAM kinases complexed with cabozantinib, these square displacements are key components to study the conformational changes in the active states of various TAM RTK family members. Axl has lower displacements compared to Mer and Tyro3 kinase when bound with cabozantinib. This indicated that Axl has greater structural stability than the other two members of subfamily. Analysing MD simulations trajectory data as vibrational modes of protein domain fluctuations gives superposition of normal modes. These normal modes can be generated from conformational fluctuations of harmonic oscillations such that all data points are represented as vectors and then converted into diagonalized matrix with respect to second order derivative of Hessian matrix.

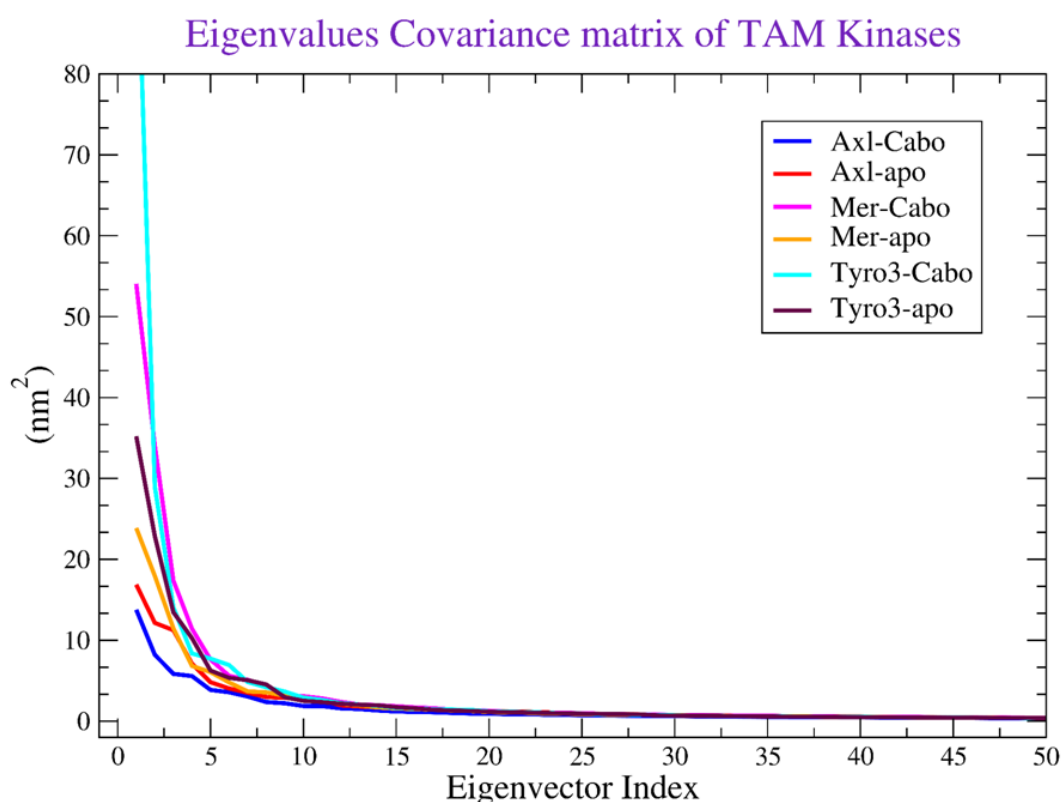


Figure 2.6) The first 50 premium Eigenvectors covariance matrix shown as cumulative sum of TAM kinase fluctuations.

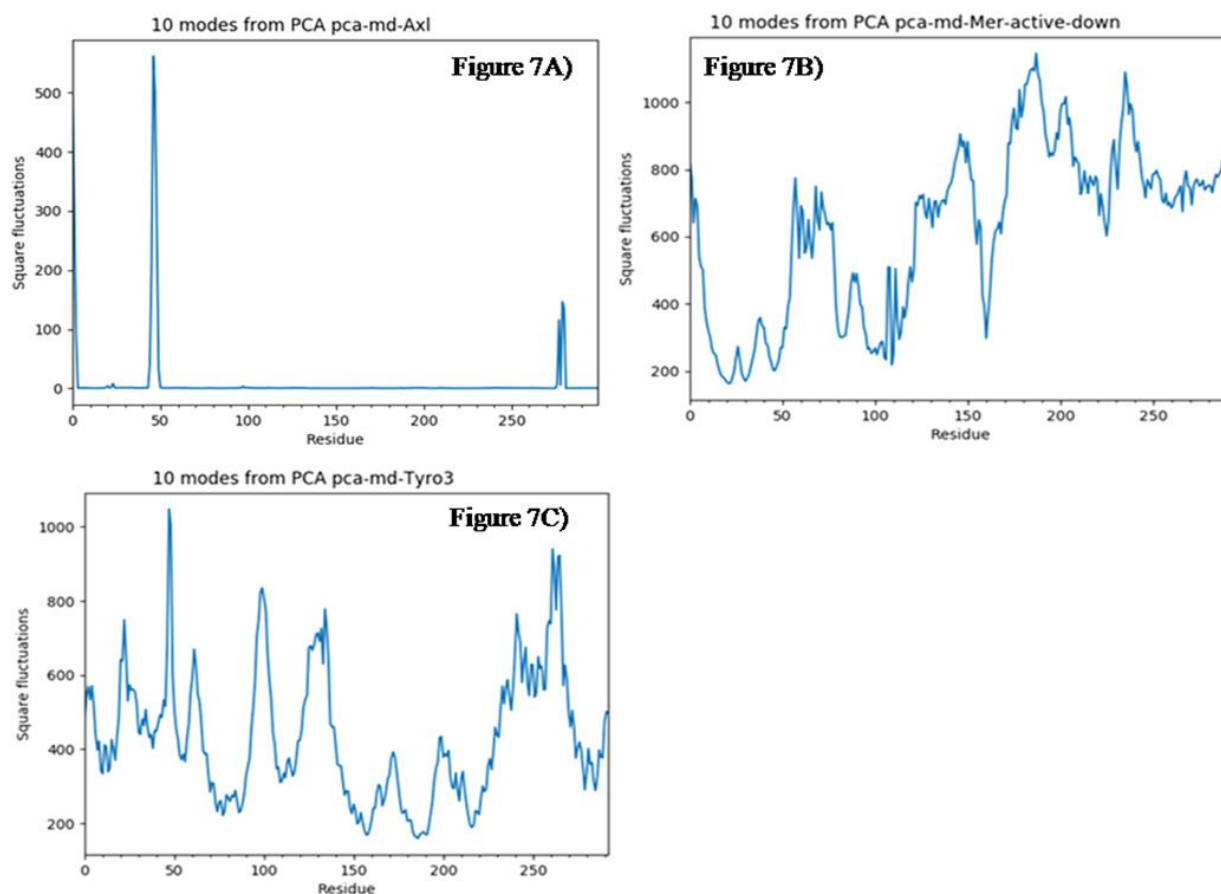


Figure 2.7) Principal component analysis of TAM kinase – cabozantinib complexes from 250 ns MD simulations. In this figure, for the sake of convenience all TAM kinases are numbered from one onwards.

The constrained motions in all simulations of protein-ligand complexes (Bakan, & Bahar, 2009). The overlap plots of ANM Vs PCA gives pre-validation of active states and populated states of kinase domain at 1 and 50 ns time scale. This is a glimpse of further data analysis with ANM of key unique motions captured for Hessian matrixised vector data to predict the active model of TAM kinase in the presence of inhibitor. As shown in Figures 2.8(A, B, C), Axl converged plot has more faint blue blocks than Tyro3, But Mer plot has more deep blue blocks which indicated lesser convergence between normal Vs. PCA modes in Mer kinase. The overlaps of ANM Vs PCA are key pictorial representation of stable active states existing in Axl compared to Mer and Tyro3 kinases. The best converged ten eigenvectors systems could be captured with the help of correlated modes between principal components versus anisotropic normal modes (ANM) of trajectories extracted from MD extrapolated in 2D coordinate system as ANM Vs. PCA generated correlated 2D plots for TAM-cabozantinib complexes before (1 ns) and after (50 ns) MD simulations time frames. “MDAnalysis” Python libraries are an

object-oriented tool for the components in a molecular system. In a protein kinase, the vibrant domain motion analysis can be performed to retrieve specified loop dynamic constraints that are crucial to explain the protein structural dynamics for “activation loop” and “ α C-helix”, in apo and cabozantinib bound TAM kinase domains. The Figures 2.10 and 2.11(A, B, C, D, E, F) clearly indicated variations in the activation loop and α C-helical dynamics in all molecular systems. The activation loop and α C-helix C α RMSD was extracted from 250 ns MD trajectory data of TAM kinases. The Axl activation loop showed greater stability or fewer conformational changes when bound to cabozantinib, compared to the apo Axl and the dynamics of the α C-helix is also greater in the apo form compared to the cabozantinib bound Axl. Similarly, apo Tyro3 has more activation loop and α C-helical dynamics compared to the cabozantinib bound form. In the apo and inhibitor bound forms of Mer RTK, activation loop has higher dynamic states indicating higher flexibility of this loop, also the α C-helix is stable when bound to inhibitor compared to the apo form. In the apo Mer kinase domain, the activation loop RMSD is highly populated and does not attain a stable conformation. In other words, the apo TAM kinases display greater activation loop dynamics compared to the inhibitor bound forms. Among the inhibitor bound states, the order of stability is Axl>Tyro3>Mer. From these plots, it is observed that in the inhibitor bound form, the TAM kinases show some periodic constraints during MD simulations that could be important for their function.

The overlap populations among the receptor-based TAM kinases can be projected to compare the initial and final average structures of proteins while interacting with ligands. The PCA correlated with the overlap plot of ANM shown in Figure 2.8 indicated that the overall populations of Axl are more than Tyro3 and Mer. However, Tyro3 has populations more than in Mer as indicated by 0.1 range cutoff, and the Axl has more populations over marginal cutoff (0.15) next to Tyro3. This is a prime indication of higher intra-residue interactions in 50 ns structures for both Axl and Tyro3 than compared to Mer-inhibitor complex. Therefore, in Axl, the number of microstate populations are more, whereas least in Mer indicating the more dynamical nature of Mer kinase domain. To further probe internal dynamics and conformational diversity in TAM kinases, general cross-correlation analysis was carried out. A complete map of the mechanical resistance in response to all possible pulling directions and the mean value of effective spring constant with secondary structure of TAM kinase domains is shown in the Figures 2.9 (A-F) and 50 ns MD trajectory data for TAM Kinases. ANM overlap modes of active micro state populations before 1 ns and 50 ns MD trajectory data for TAM kinase fluctuations (Cutoff range 0.1).

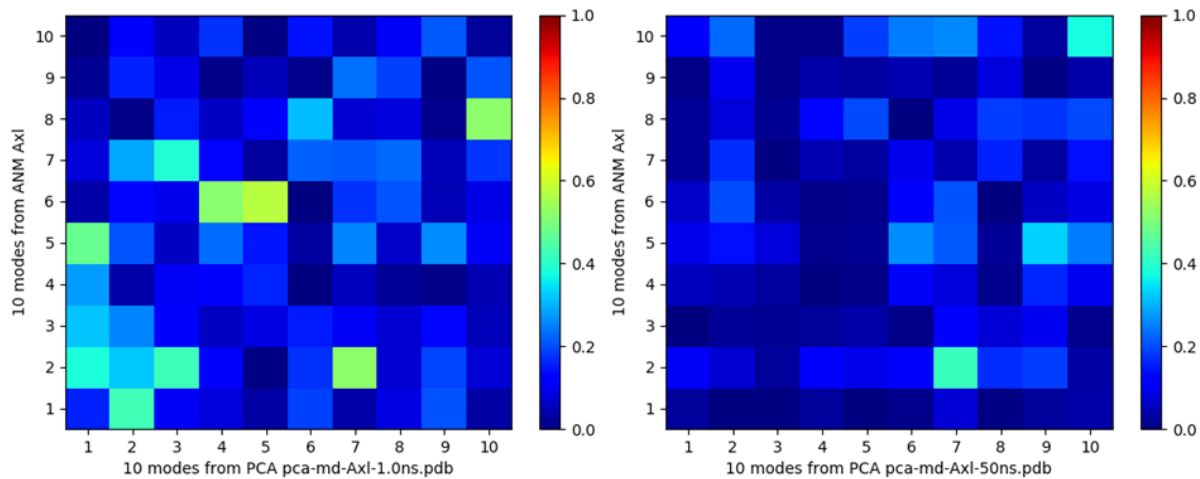


Figure 8A)

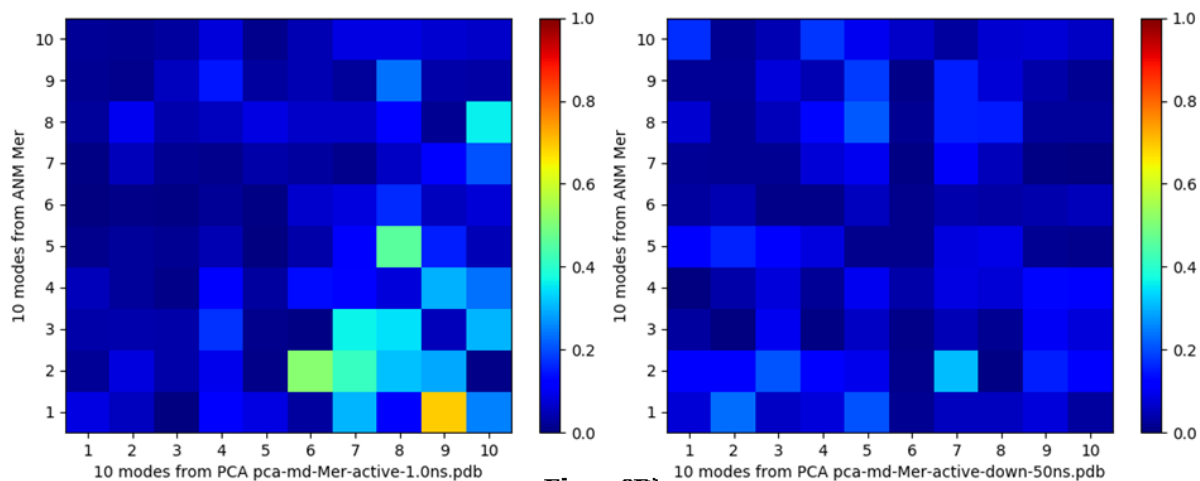


Figure 8B)

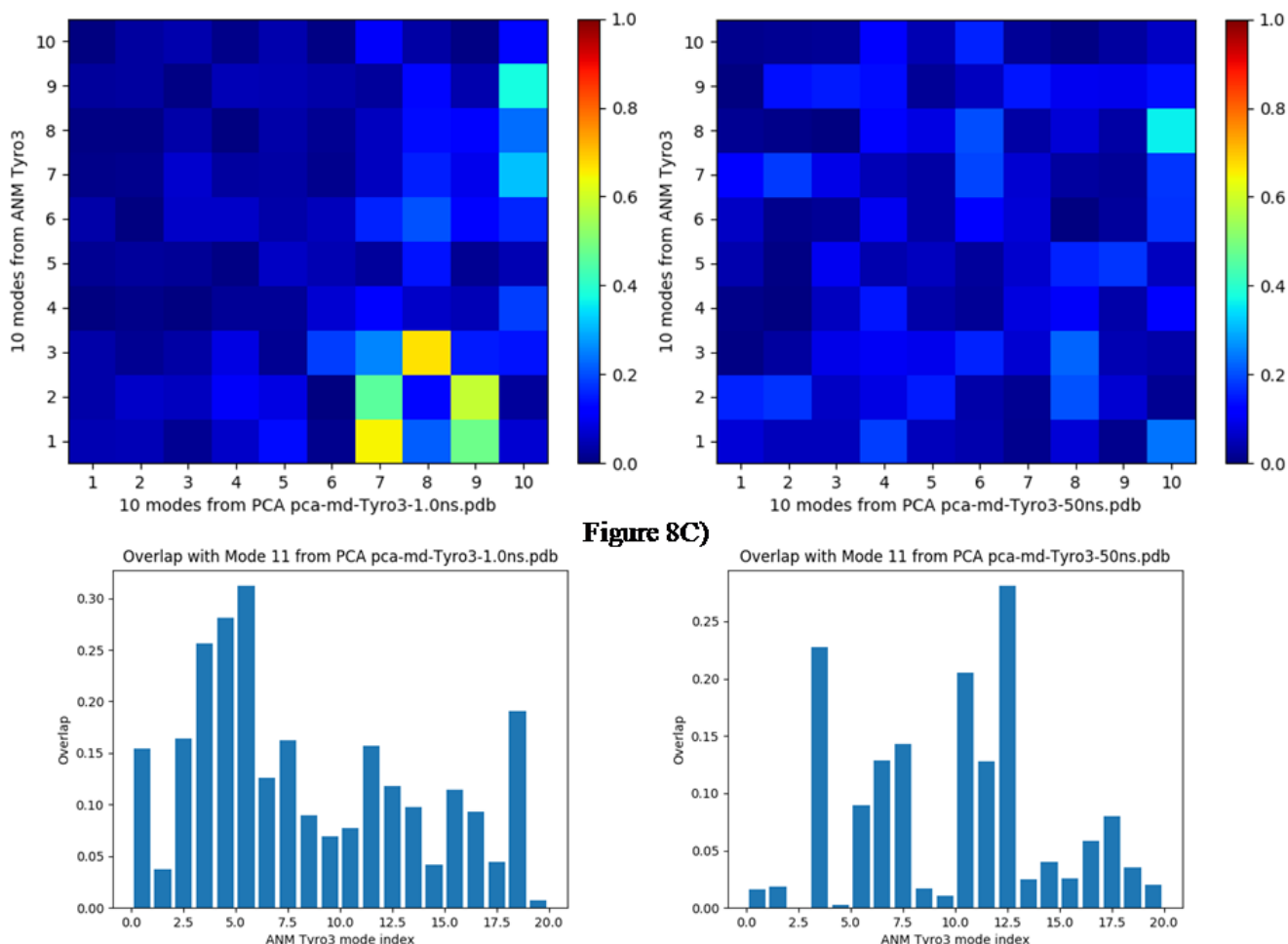


Figure 2.8 (A, B, C) Cross correlation (faint Blue blocks) between ANM Vs PCA before 1 ns.

The correlated extents of protein motion are different in all molecular systems. Certain regions in the protein show strong correlated motion while other regions display low correlated motion. Apo and cabozantinib bound Axl kinase display similar regions and extent of cross correlations. In comparison, the apo Mer displays lower correlated motion compared to the inhibitor bound form. The apo and inhibitor bound Tyro3 exhibit different extents of cross correlations in different regions of the proteins. The correlated plots also explain the distortedness and stiffness of protein in the presence of inhibitor bound in the active site of TAM RTKs. These plots have cross correlation of C α residues in the proteins from same subfamily which describes the structural deformations of protein in the absence and presence of high affinity inhibitor in the active site. In comparison, the Axl kinase has more mechanical stiffness indicating its structural rigidity among them.

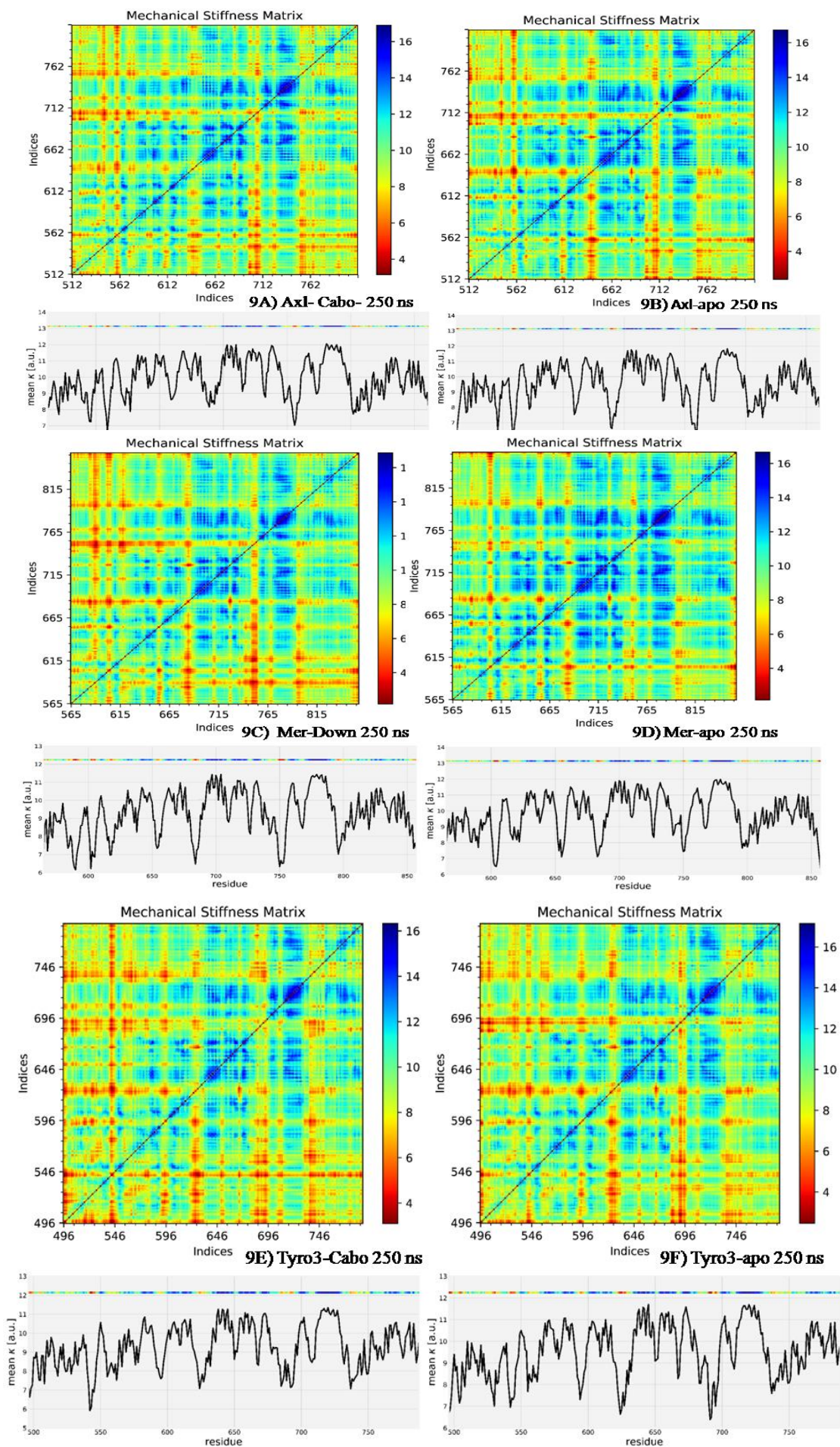


Figure 2. 9 (A, B, C, D, E, F) The complete mechanical stiffness 2D plot of TAM kinases (blue mechanically strong regions, red - mechanically weak regions). Color bar indicates effective force constants. The mean value of effective spring constant and the protein secondary structure is shown as 1D profile.

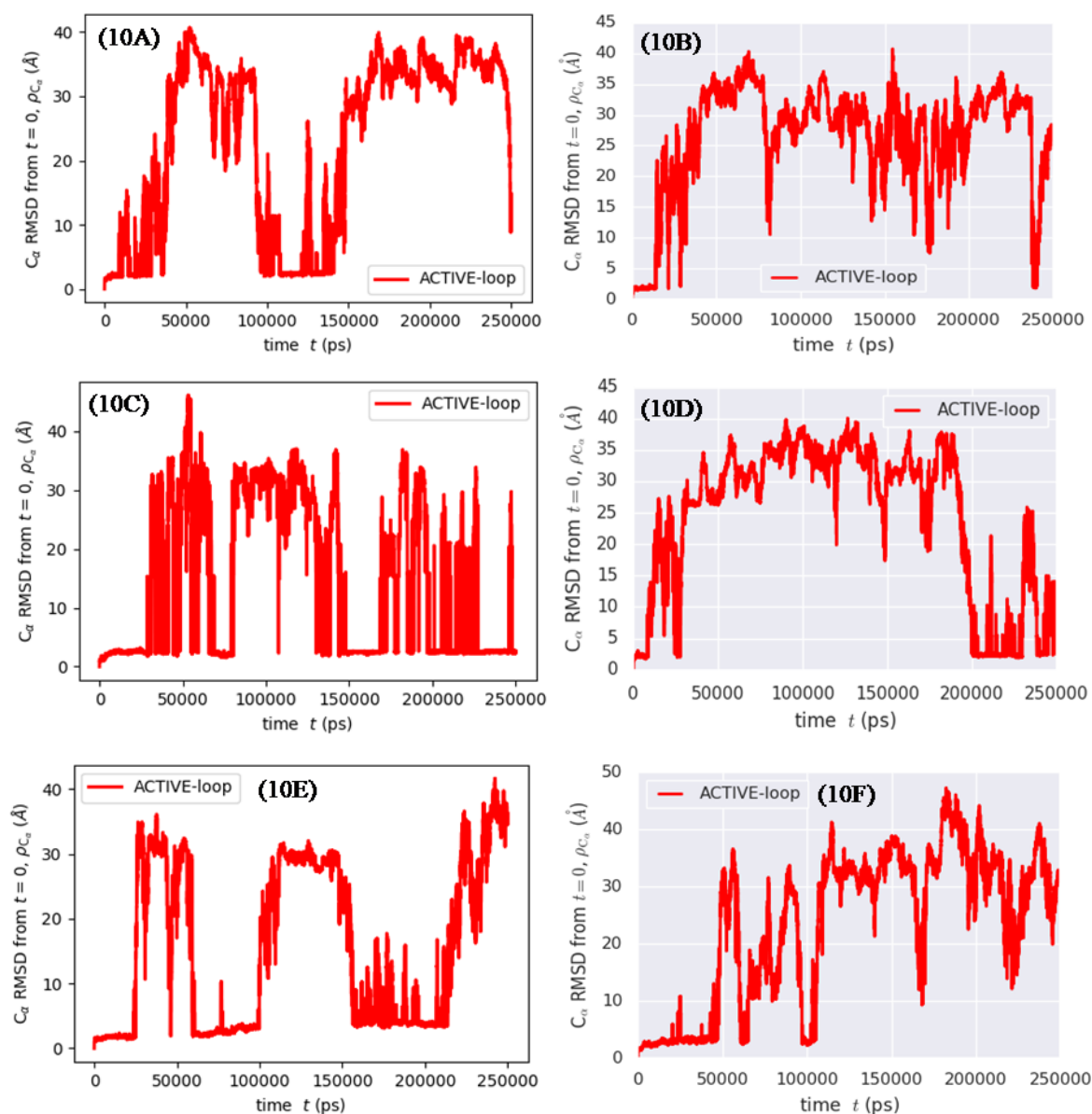


Figure 2.10) MDTraj Analysis of C α RMSD of activation loop dynamics in TAM RTK's cabozantinib and apo models at 250 ns. In Figure: Active Loop dynamics of 2.10A) Axl-inhibitor; 2.10B) Axl-apo. In Figure: Active Loop dynamics of 2.10C) Mer-inhibitor; 2.10D) Mer-apo. In Figure: Active Loop dynamics of 2.10E) Tyro3-inhibitor; 2.10F) Tyro3-apo

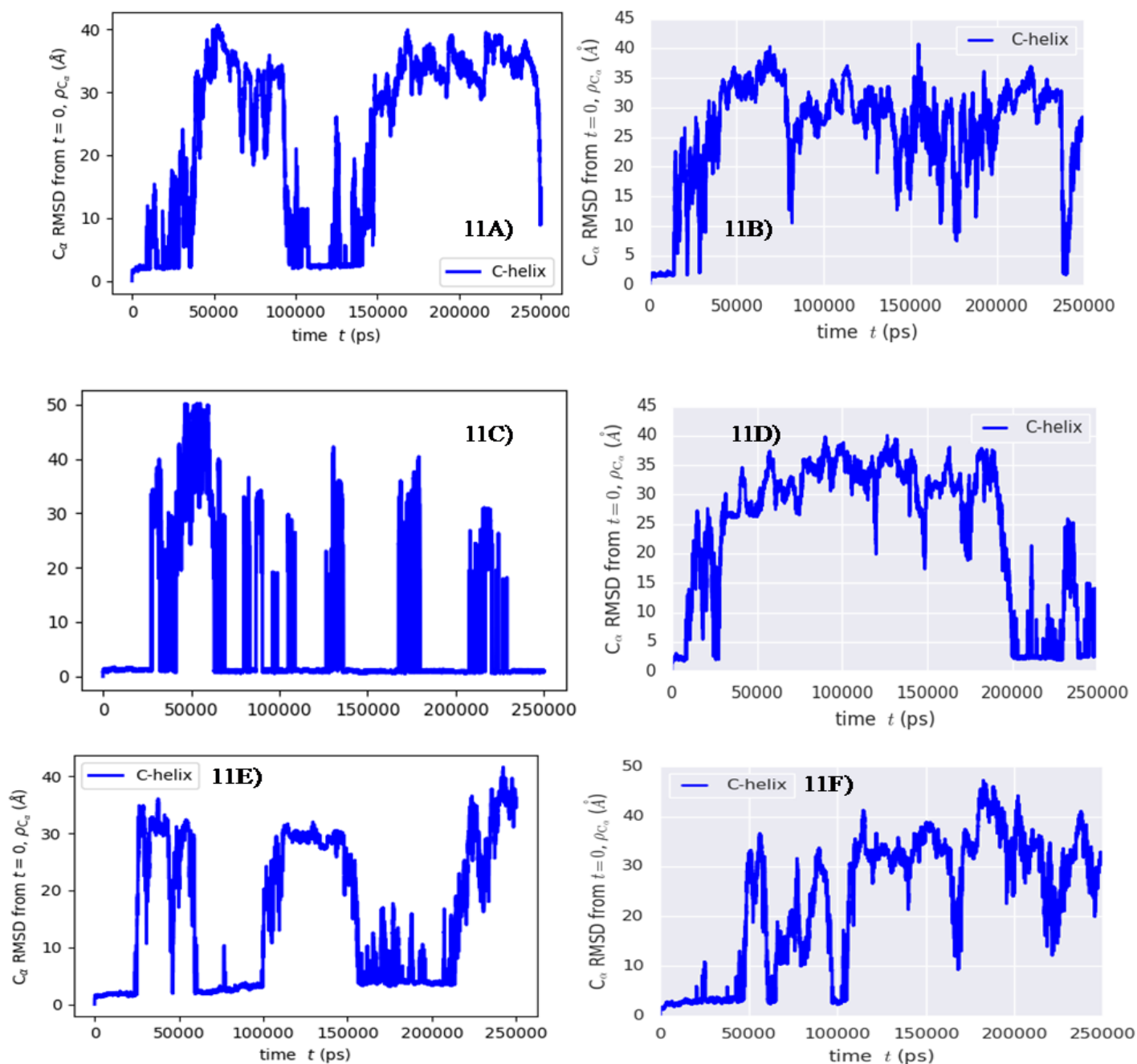


Figure 2.11) MDTraj Analysis of $C\alpha$ RMSD of α C-helix dynamics in TAM RTK's – Cabozantinib and apo models at 250 ns.

In Figure: α C-helix dynamics of 2.11a) Axl-inhibitor; 2.11b) Axl-apo

In Figure: α C-helix dynamics of 2.11c) Mer-inhibitor; 2.11d) Mer-apo

In Figure: α C-helix dynamics of 2.11e) Tyro3-inhibitor; 2.11f) Tyro3-apo

The α C-helix of Mer receptor kinase bound to cabozantinib has higher C α RMSD compared to Axl and Tyro3. The Mer activation loop dynamics shows the possible existence of metastable active forms among the TAM RTKs. The MSM prediction is studied with Python coding and specialized scientific Bio-python, interactive python environment in the Linux platform. The MSM building carries out featurisation, clustering, projection and estimation that was done with pyEMMA 2.5.5 Python package. Measurement of the time dependent C α distance variations in the TAM kinases is a strategy for Markov model interpretation. The 250 ns MD simulations trajectory data is utilized for Markov's model building and is shown in Figure 2.12 (A, B, C, D, E, F).

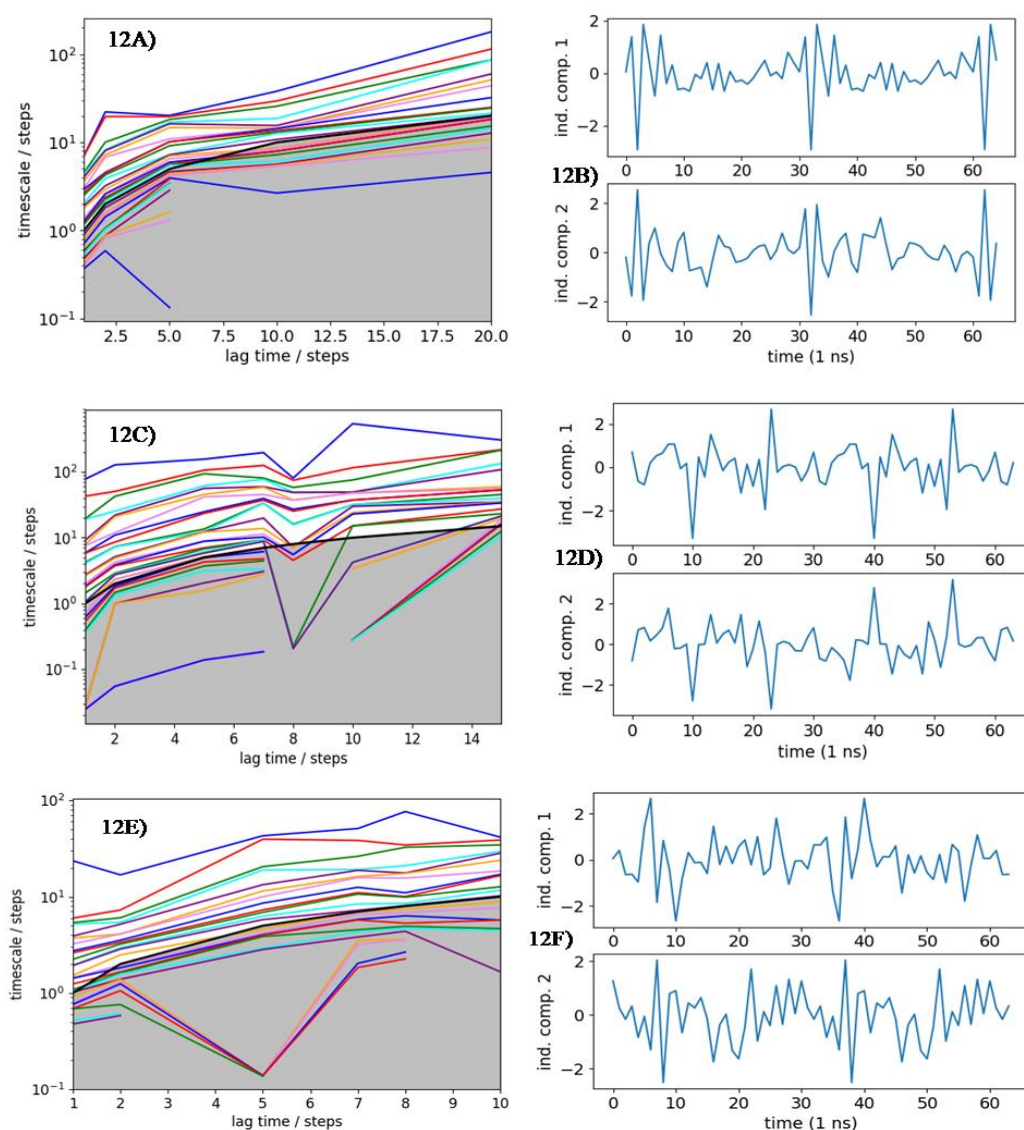


Figure 2.12) MSM models and tICA analysis of C α RMSD TAM RTK's – cabozantinib complexes from 250 ns.

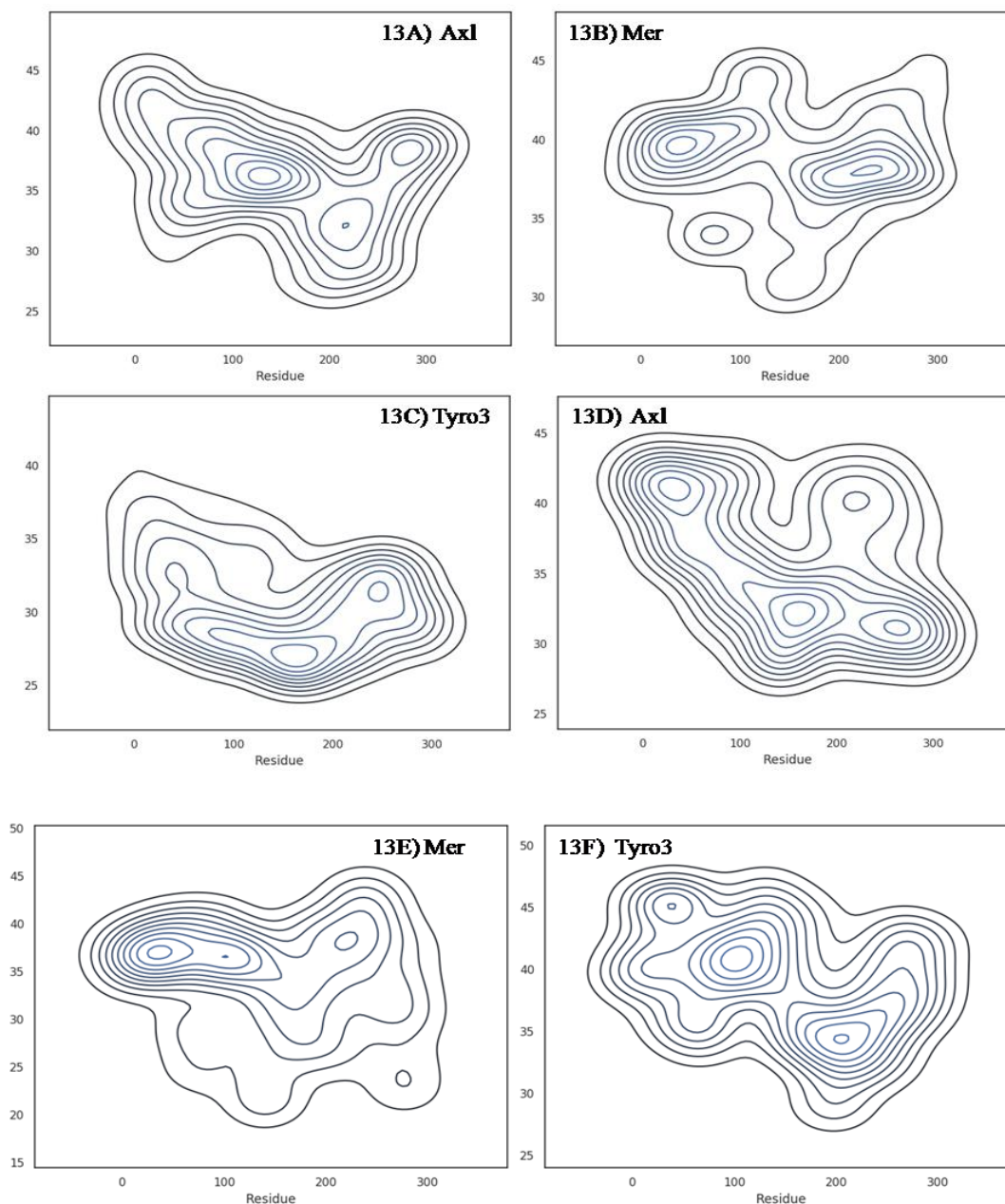


Figure 2.13 A, B, C) MDTraj analysis of RMSF contour plot TAM RTK's – Cabozantinib - forms 250 ns, (all Y axis units Å).

Figure 2.13 D, E, F): MDTraj analysis of RMSF contour plot TAM RTK's – apo- forms from 250 ns.

Residue Interaction Networks

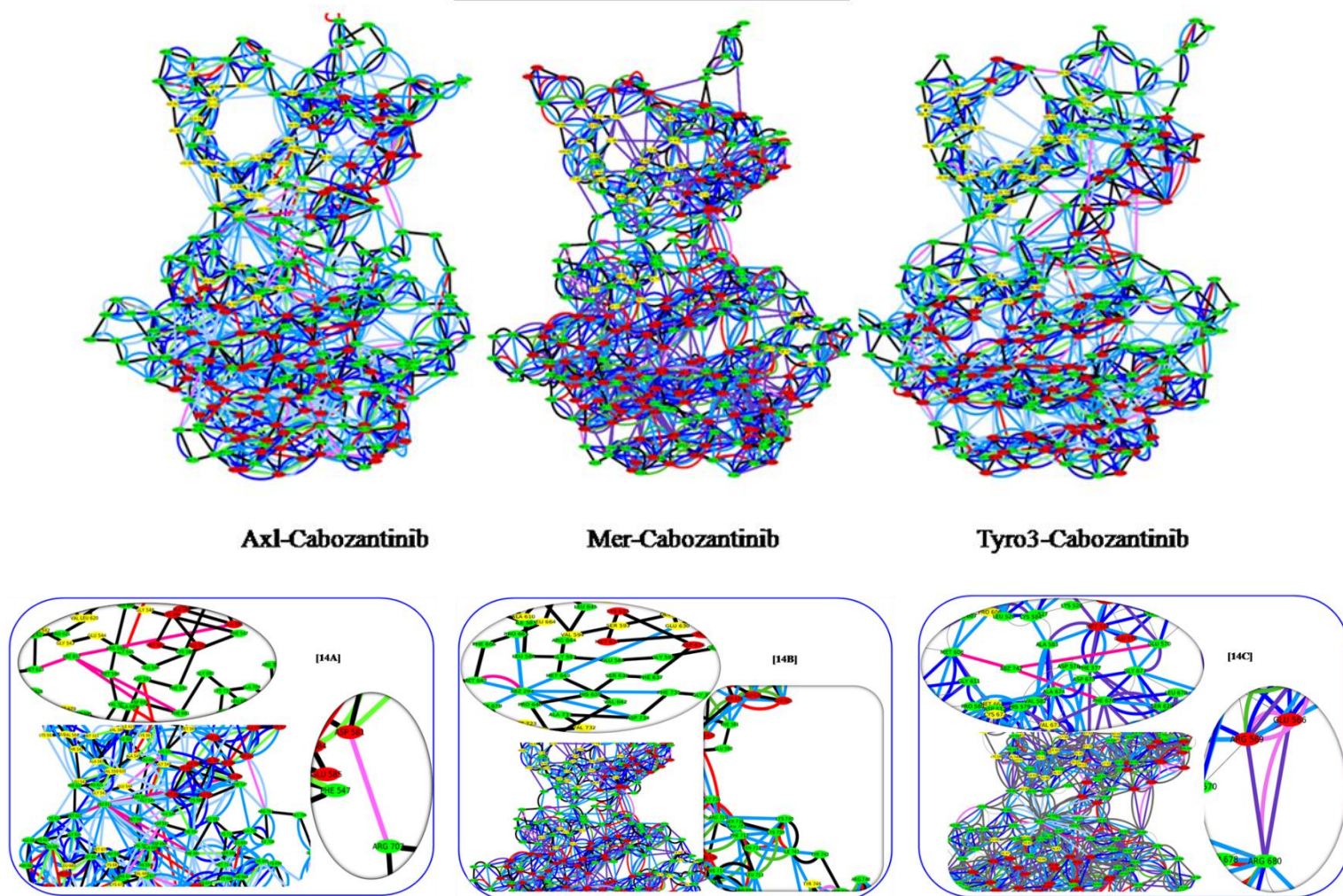


Figure 2.14A) Axl-cabozantinib-RIN plot. It has intra chain interaction (Asp-581 – Arg-707).

Figure 2.14B) Mer-cabozantinib-RIN plot. It has lack of intra chain interaction (Glu-626 – Arg-748).

Figure 2.14C) Tyro3-cabozantinib RIN plot. It has intra chain interaction (Glu-566 – Arg-680).

The MSM building and populated active states are shown to vary in Axl (specified populated), Mer (discrete populated) and Tyro3 has intermediate populated states. MSM model prediction is in specified time intervals (lag time = 35 out of 64) of 1 ns data set of state-space discretization. The populated models can be deducted with dimensionality reduction from 879 into 7 by using tICA which can be projected in the form of a kinetic map as indicated in the Figure 2.12. The independent components vs time interval (1 ns) from the recorded trajectory after every 0.5 ns frame up to 250 ns of MD simulations data are shown in this plot. In Mer-cabozantinib bound complex, periodic constraints of C α distance indicates that it should

comprise more metastable Markov's model populations among them. The RMSF contour plots (Figure 2.13) were sketched with help of scikit-learn Python libraries. Mer RMSF contour plot displayed some key radiant contours of specified states visualized in cabozantinib bound Mer kinase. In Axl and Tyro3- cabozantinib bound forms, the active states are well mixed up. However, in the apo form, the active states are more mixed up and even indistinguishable in apo Mer contour, whereas the apo Axl and Tyro3 have discretized specific contours. In addition, the RMSF contours of apo TAM kinases have high oriented specific residue α motions observed but the cabozantinib bound receptor kinase displays low RMSF of specified residue motions given the residue index on X-axis. The inhibitor bound Mer has well distinguished contour state plot. Based on the above observations, Mer appears to be a special class of kinase due to the existence of more active metastable states compared to Tyro3 and Axl. The large domain motions occurring in Mer are due to the loss of intra-chain hydrogen bond between Glu626 – Arg748 during the course of MD simulations (Figures 2.14b). In the case of Axl, this interaction is present farther away from the active site in Axl-inhibitor binding region, therefore this could tightly hold N-terminal domain in diagonal manner (mediated via main chain-side chain non-bonding interactions) in Axl-cabozantinib complex. Tyro3 has this corresponding intra-chain interaction very close to the active site of the protein, therefore it does not restrict full domain motion when compared to Axl– cabozantinib complex. These non-bonding interactions between residues in the N- and C-terminal domains play a significant role in the constrained motions of TAM kinases in the presence of suitable ligand environment (cabozantinib). Intra-chain interactions could be key factors responsible for constrained motions in protein, but protein-ligand complex formation is ultimately an indirect way to orient these intra-chain non-bonded contacts by virtue of protein dynamics and internal motion. The RIN plots showed key interactions among TAM kinase-cabozantinib complexes to satisfactorily explain that dynamic motion in proteins. It is crucial to explain variable domain motion constraints occurring due to the hyper sensitivity of non-bonded interactions among various residues within the protein domains. The intra-chain interactions present in Axl (Asp581 – Arg707) (Figure 2.14A) and Tyro3 (Glu566 – Arg680) (Figure 2.14C) have dominant role in constrained motions of protein-inhibitor complexes which confirms the changes in structural features of protein. The domain motion frames are captured through α atomic NMA using VMD software that showed vibrionic motions between inter-domains in selective protein motifs. It was observed that the Axl normal modes oscillated as stretching mode, whereas, Tyro3 vibrated in intermediate rocking mode as compared with Mer that has super rocking modes (Figure 2.15 A, B, C). This is an essential observation of domain motions

in the subfamily of kinases in the presence of binding inhibitors essentially due to specific non-bonding interactions among various residue motifs and therefore the variations in the structural plasticity of TAM RTKs.

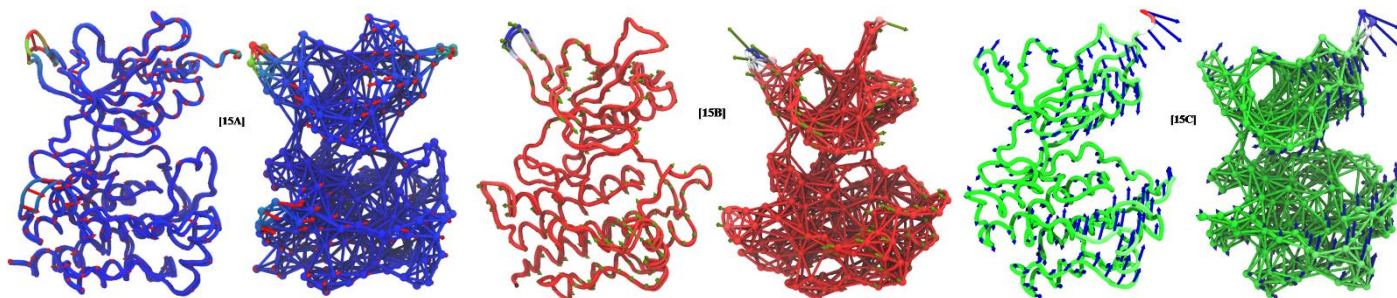


Figure 2.15A) Porcupine plot Axl-cabozantinib-250ns (ANM modes).

2.15B) Porcupine plot Mer-cabozantinib-250ns (ANM modes).

2.15C) Porcupine plot Tyro3-cabozantinib-250ns (ANM modes).

2.4. Conclusions

Tyro3, Axl and Mer are TAM RTKs that belong to the family of integral membrane proteins and share high sequence similarities. TAM kinases are key receptors for discovery of effective and specific cancer drugs that bind to either active or allosteric sites using methods in structure based drug design. Comparison of the crystal and models of the active states in TAM kinases revealed their high structural similarities. Cabozantinib, a nanomolar affinity TAM kinase inhibitor binds to the ATP binding pocket of the enzymes and is stabilized by several non-bonding interactions. The structural stabilities of apo and cabozantinib bound TAM kinases were studied using 250 ns MD simulations, the RMSD plots revealed the stability of the molecular systems and the RMSF plots indicated the fluctuations in the N-terminal domain motifs and the activation loop. An in-depth study of TAM kinase MD trajectory analyses revealed that the non-bonding interactions mediated via three residue contact involving the $\beta 3$ strand, αC -helix and DFG motif decide the retention or expansion of the active core of the kinases. Further, the ionic interactions between oppositely charged residues on αC -helix (acidic) and the activation loop (basic) is responsible for major domain motions in TAM kinases. The fluctuations in the αC -helix and activation loop regions lead to the presence of diverse conformations in Mer and Tyro3. Among the TAM kinases, the binding stability of Tyro3 with inhibitor is intermediate between Axl and Mer. This work revealed the active metastable states in TAM kinases and the molecular mechanism of function regulation. This classical MD simulations study on apo and TAM kinases-cabozantinib complexes gives

description of the non-bonding interactions at specified sites in the protein active cavity. These atomistic simulations and analyses revealed the existence active microstates in TAM kinases and further strengthen the overwhelming literature on the dynamic states in kinase family proteins.

Chapter -3

Dynamic conformational states of apo, ATP and cabozantinib bound TAM kinases to differentiate active-inactive kinetic models

Abstract

The dynamically active and inactive conformations of kinases play a crucial role in the activation of intracellular downstream signaling pathways. The all-atom MD simulations at microsecond (μ s) timescale and longer provide robust insights into the structural details of conformational alterations in kinases that contribute to their cellular metabolic activities and signaling pathways. TAM RTKs are overexpressed in several types of human cancers. Cabozantinib, a small molecule inhibitor constrains the activity of TAM kinases at nanomolar concentrations. The apo, complexes of ATP (active state) and cabozantinib (active and inactive states) with TAM RTKs were studied by one μ s MD simulations followed by trajectory analyses. The dynamic mechanistic pathways intrinsic to the kinase activity and protein conformational landscape in the cabozantinib bound TAM kinases are revealed due to the alterations in the P-loop, α C-helix and activation loop that result in breaking the regulatory and catalytic spines, while the active states of ATP bound TAM kinases are retained. The co-existence of dynamical states when bound to cabozantinib was observed and the long-lived kinetic transition states of distinct active and inactive structural models were deciphered from MD simulations trajectories that have not been revealed so far.

3.1. Introduction

The RTKs are single-pass membrane spanning proteins, recognised by specific extracellular ligands to cause receptor dimerization followed by kinase activation and intracellular autophosphorylation (Endicott et al., 2012). TAM RTKs are overexpressed in acute myeloid leukaemia, breast, colorectal, lung, ovarian cancers and glioblastoma. These three TAM kinase members share a high degree of sequence and structural homology in their kinase domains. Cabozantinib was approved by FDA for advanced renal cell carcinoma, hepatocellular carcinoma and medullary thyroid cancer and also differentiated thyroid cancer that has progressed following prior VEGFR-targeted therapy. Cabozantinib is reported to bind TAM kinases with high affinity at nanomolar concentrations (Gajiwala et al., 2017, Sultan et al., 2017, Skora et al., 2013, Turner and Blythe, 2019, Herum, et al., 2017, Qin et al., 2019, Lacy et al., 2018, Pantano et al., 2016, Myers et al., 2019, Robinson, 2013). The distinction between the active and inactive states in a kinase is based upon the α C-helical movement towards or away from the ATP binding site. The presence of catalytically important Lys567 (close to P-loop) - Glu585 (α C-helix) is an essential ionic interaction in the active Axl kinase from the crystal structure in PDB. The disordered activation loop (689-724) in the C-terminal lobe has altered conformational states that are variable among the kinase structures reported so far. An ionic interaction between the side chains of Asp581 (α C-helix) and Lys695 (activation loop) is important in the kinase structure and allostery. The synchronous fluctuations in the P-loop, α C-helix and activation loop leads to spatial alteration in the shape of the enzyme active site pocket and distinct structural features such as the inward/outward rotation of α C-helix and expansion of the activation loop. The Lys567 - Glu585 salt bridge is the indication for the active state of TAM RTKs. A kinase domain has two kinds of active sites; regulatory substrate site and catalytic active site that become available during allosteric competitive inhibitor binding pathways in the cellular signal transduction process. Structure analyses revealed the presence of two non-contiguous structural motifs termed regulatory and catalytic spines (Robinson, 2013, Mohanty et al., 2016, Hu et al., 2015) that are required for stabilizing the protein in the active state. Since the Axl kinase domain is crystallized in both active and inactive forms, classical long range MD simulations were performed on the active and inactive states of TAM kinases to obtain key insights into the spatial dynamics and to understand the cellular mechanistic pathways of inhibitor, cabozantinib binding to kinases that will prevent internal signaling by up-regulation or overexpression of kinases. In this chapter, the highly unstable conformational transition states including regulatory and catalytic spines in the kinase domains are reported by studying the apo, ATP and cabozantinib bound TAM RTKs each for 1 μ s MD simulations using AMBER 18.14 suite of programs.

3.2. Materials and methods

3.2.1. Structures of apo, active and inactive TAM RTK kinase domains

The three-dimensional crystal structures of Axl (PDB id: 5U6B) (Gajiwala et al., 2017) A and B chains exist as inactive and active states, respectively. The missing residues in the activation loop were built using "Model/Refine Loops" in "Structure Editing" tool in UCSF Chimera 1.12. (Yang et al., 2012) The active and inactive homology model structures of Mer and Tyro3 were built based on the crystal structures of 5U6B, B and A chains, respectively, using MODELLER (Sali and Blundell 1993, Webb, and Sali, 2014) as described previously.

3.2.2. Molecular docking of ATP and cabozantinib

The inhibitor cabozantinib was docked into the ATP binding pocket of the active and inactive conformers of TAM kinases and ATP was docked into the ATP binding pocket of the active state of TAM kinases using AutoDock (Morris et al., 2009). The docking pose with lowest binding energy and maximum docking poses was utilized for further MD simulations to decipher the molecular basis for interactions between protein and ligand.

3.2.3. Molecular dynamic simulations

All MD simulations were achieved using AMBER (Gotz, et al., 2012) version 18.14 for the apo, ATP bound active, and cabozantinib bound active and inactive states of TAM kinases. The best docking pose of each complex was utilized as input for MD simulations. The force fields for the entire systems were generated with Antechamber using am1bcc method (Wang et al., 2006, Colovos and Yeates, 1993). All input parameter files for MD simulations were generated after adding hydrogen atoms in tLEaP module in AMBER tools (Anandakrishnan, Aguilar and Onufriev., 2012, Lindorff-Larsen et al., 2010). Sodium and chloride ions were added to the systems to neutralize the charge, each molecular system was solvated within a 10 Å size box. The final ionic concentration for the systems was set to 100 mM. The Amberff99sb-idln force field was used for entire model system with TIP3P water model for AMBER molecular parameters (Meagher et al., 2003, Mark and Nilsson, 2001). All MD simulations were run at 300 K temperature and 1 atm pressure with Monte Carlo barostat (Salomon-Ferrer et al., 2013). Energy minimization was carried out by using steepest descent method for 40,000 cycles to overcome short range null contacts among the molecular system in solvent (Darden et al., 1993). Long range electrostatic interactions were considered with Particle Mesh Ewald algorithm (Jorgensen et al., 1983) with cut-off range 9 Å and order 4. All model systems were equilibrated for 5 ns before the production run, and the coordinates in the production run were saved after every 5 ps (Salomon-Ferrer et al., 2013, McGibbon et al., 2015). The MD simulations of each molecular system was carried out for 1 μs, accounting for a total of 12 μs simulations time.

3.2.4. Data analysis

The MD trajectory data analysis was carried out using cpptraj and pytraj in Amber tools 18 (Hornak et al., 2006). The average structures after MD simulations, RMSD, RMSF and PCA were derived from the trajectory analysis. For the sake of data space minimization during post MD analysis, the Markov state model (MSM) analysis was carried out on 40K frames out of 200K frames and the PCA was carried out on the data from 1K frames generated from each molecular system. To build the MSM, datasets of close accessible kinetic metastable states associated with protein conformational ensemble obtained from large scale simulations are required. These states can be defined in pyEMMA Python library (Scherer et al., 2015). To generate the MSMs 40K conformations were sampled. All twelve MD simulations datasets were transformed in terms of protein C α backbone dihedrals, C α backbone atomic positions and distances from their trajectories. All MD simulations trajectories were analysed for 1000 ns (200K frame data) by sampling the MSM predictions (Harrigan et al., 2017). This identified kinetically metastable transitions among cluster k-means lag time (250 degrees of freedom) of protein conformations (Perez-Hernandez et al., 2013). The extrapolation of the real time data into pictorial and graphic vectorized data points was achieved with matplotlib and numpy data frames into 2D plotting space. The state distributions of kinetic metastable data points were featurized and cluster analysis was applied using TICA (Perez and Granger, 2007, Pedregosa et al., 2011, Noe and Clementi, 2015).

3.3. Results and discussion

The amino acid sequence alignment of Tyro3, Axl and Mer kinases shown as the final modeled (after MD simulations) structures of the active and inactive kinases display significant conformational alterations in the P-loop, α C-helix and activation loop as shown in the Table 3.1A. The three-dimensional structures of active and inactive forms of Axl kinase domain were taken from the crystal structure (5U6B) (Gajiwala et al., 2017) B and A chains, respectively. The homology models of active and inactive forms of Mer and Tyro3 kinase domains were constructed and validated. The models of TAM kinase domains constructed were compared with model structures generated using Phyre2 (Kelley et al., 2015) and AlphaFold (Varadi et al., 2021) by structure superposition. the structures superpose with low RMSD. From the docking of cabozantinib into TAM kinases, it was observed that it binds to the ATP binding pocket mediated via several non-bonding interactions. The hinge region residues Phe622, Met623 (Axl kinase domain) interact with dimethoxy quinoline ring nitrogen of cabozantinib. The para-fluoro phenyl interacts with Phe691 aromatic ring (DFG motif in Axl) and Asp690 forms hydrogen bond with amide nitrogen located between the cyclopropyl and phenyl rings of the inhibitor. The cofactor ATP binds the active site of TAM kinases, intermolecular hydrogen bonding interactions are observed with Pro621 (hinge region) and Asp627 (hinge region), Asn677 (catalytic loop region) in the Axl kinase domain. The

structures of apo, active TAM kinases complexed with ATP, active and inactive TAM kinases complexed with cabozantinib were subjected to 1 μ s MD simulations each, using AMBER18. Throughout MD simulations all molecular system appeared to be stable as observed from temperature vs time and total energy vs time plots

3.3.1. Active – inactive kinetic state models of TAM RTKs

From the long range MD simulations of TAM RTKs the kinetic state models are defined according to the internal structural dynamical features such as P-loop (544–549 amino acid residues), α C-helix (576 – 591) and activation loop (689–724) from the trajectories of the MD simulations data. The active/inactive conformers of TAM kinases are clearly distinguished. In the active state, the side chain of Glu585 on α C-helix is rotated inwards towards the substrate binding site to make salt bridge interaction with Lys567 in the case of ATP bound Axl RTKs. The side chain of Asp690 from the DFG motif also projects towards the active site. The outward orientation of Glu585 side chain away from the substrate to dissociate the ionic interaction with Lys567 (P-loop), and rotation of Asp690 side chain inwards into kinase active site is indicative of an inactive state of kinase (Gajiwala et al., 2017). In the inactive state, the α C-helix undergoes outward rotation, followed by the activation loop inward folding to minimize the drug binding active site that can be seen from Figure 3.1A. These are the key structural features implicated in the regulation of protein kinase activity and influence the effective binding of inhibitors. The binding of cabozantinib influenced various states of active/inactive models in Tyro3, Axl and Mer kinase domains. The active kinetic models are indicated by the ionic interaction between Lys567 and Glu585, inward rotation and activation loop extended to further maximize inhibitor binding site. In the ATP bound active TAM kinase structures, this ionic interaction is retained throughout the MD simulations (Figure 3.1B) indicating that ATP bound TAM kinases retain the active state. The catalytic spine and regulatory spine dictate the positions of ATP and substrate in the kinase domain. These spines play a key role in the catalysis of kinases while binding with ATP. The locations of regulatory spine and catalytic spine on the structures of TAM kinases based on the structures of C-Src (Robinson, 2013). The regulatory spine consists of four non-consecutive hydrophobic amino acid residues aligned vertically from N-terminal lobe towards the C-terminal lobe through the activation loop (Kim et al., 2017). These hydrophobic residues in Axl kinase domain are Leu600 (β_4 -strand); Met589 (α C-helix); Phe691 (DFG motif); His670 (catalytic loop) and an additional residue Asp731 from the C-terminal lobe (Figure 3.2 A, B, C). The catalytic spine consists of eight non-consecutive hydrophobic amino acid residues aligned vertically from N-terminal lobe towards the C-terminal lobe through the hinge region. These hydrophobic residues in Axl kinase are, Val550 (β_2 -strand), Ala565 (β_3 -strand), Phe622, Leu628 (hinge region), Met679, Leu680 (catalytic loop), Met739, Ile742 (α F-helix from C-terminal lobe).

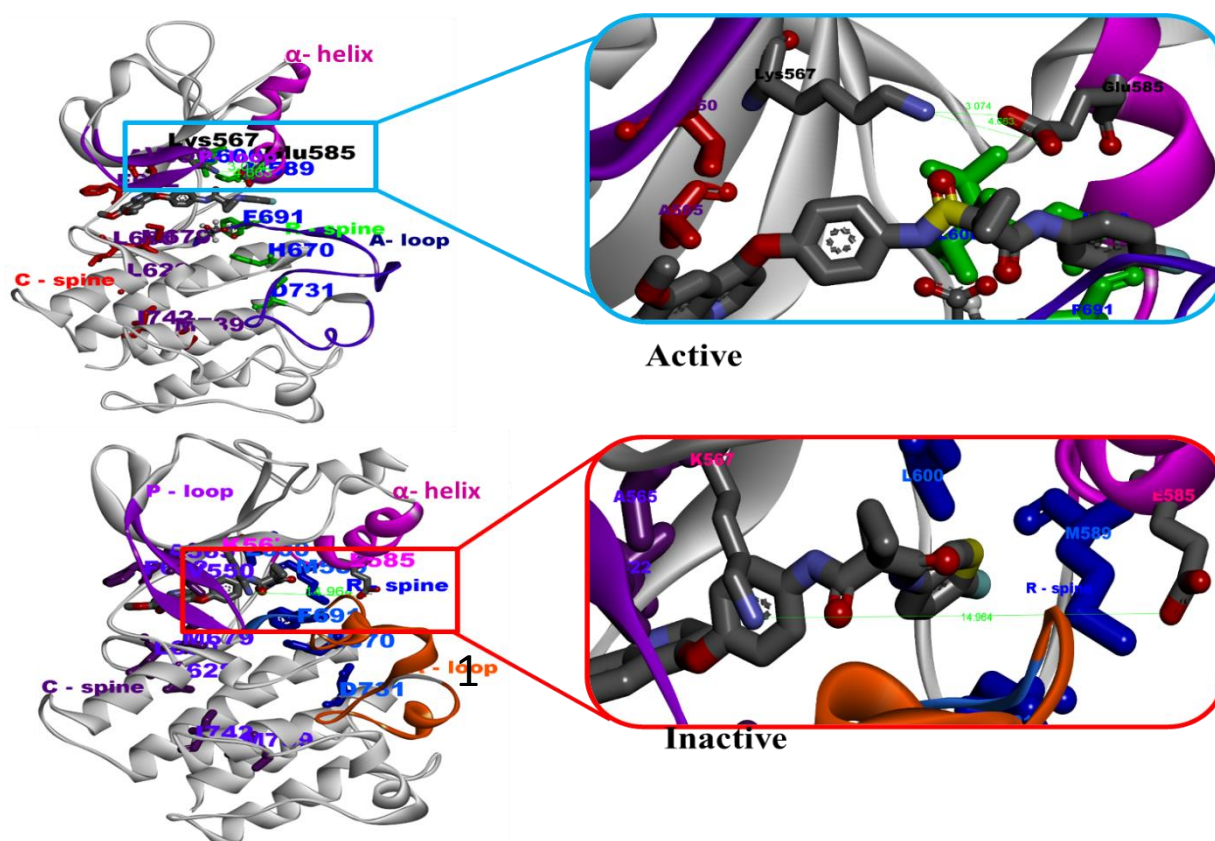


Figure 3.1A) The active and inactive states of cabozantinib bound Axl kinase domain after 1 μ s MD simulations. Axl (grey); cabozantinib (elemental color). Lys567 - P-loop with Glu585 - α C-helix.

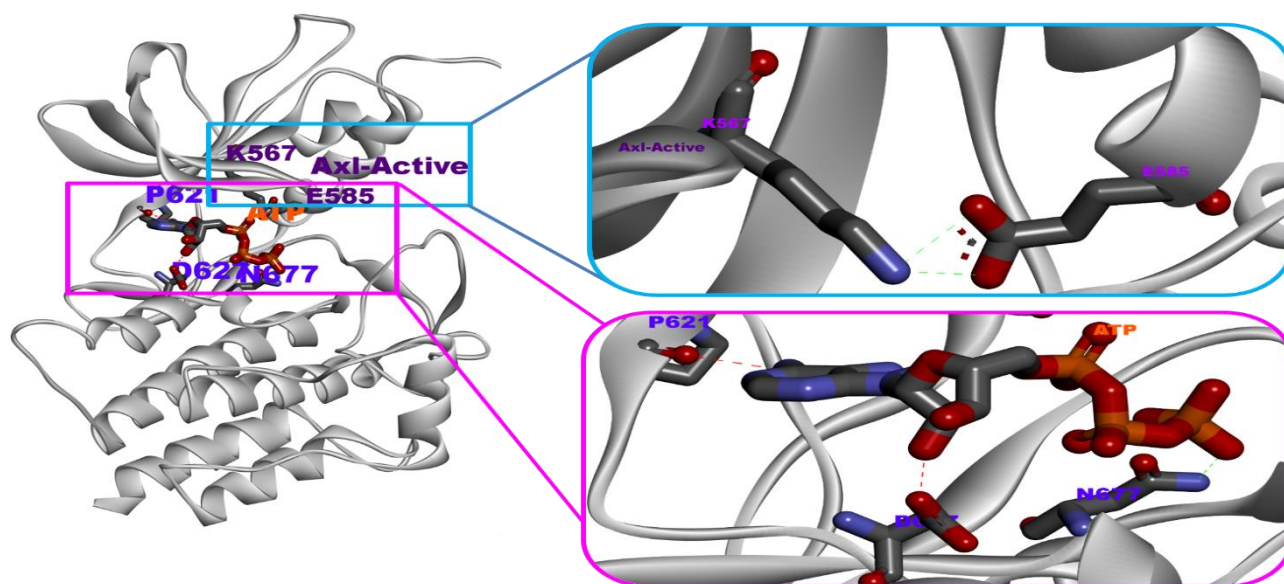


Figure 3.1B) The active states of ATP bound Axl kinase domain. Axl-active (grey); ATP (elemental color). Lys567- P-loop - Glu585- α C-helix.

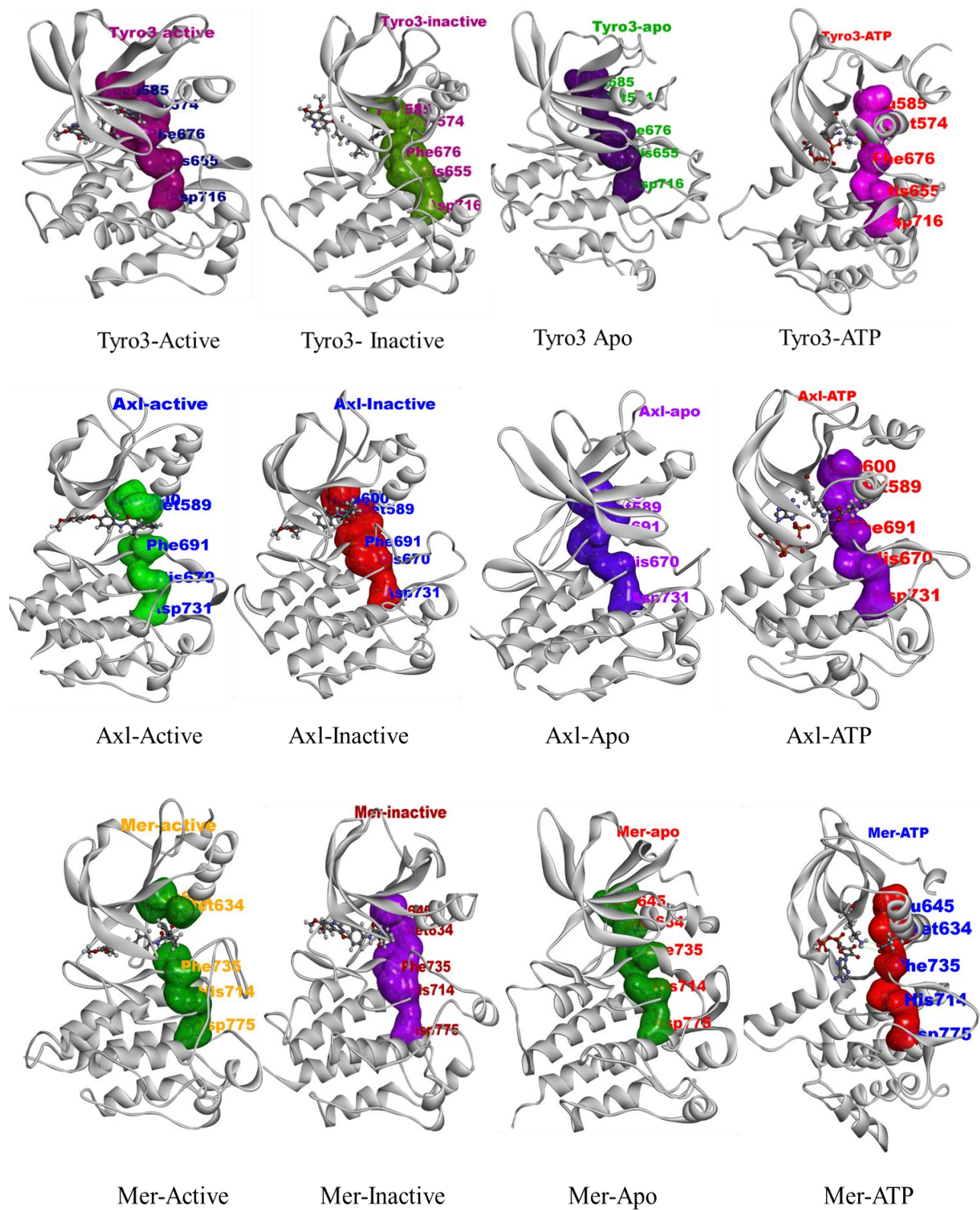
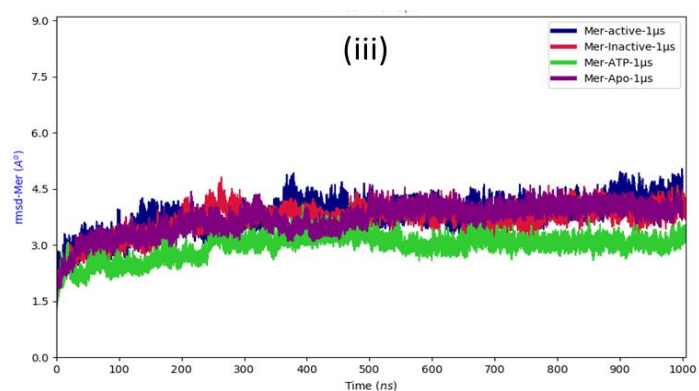
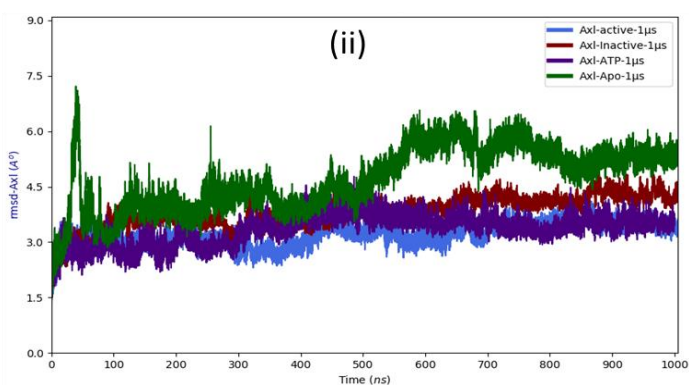
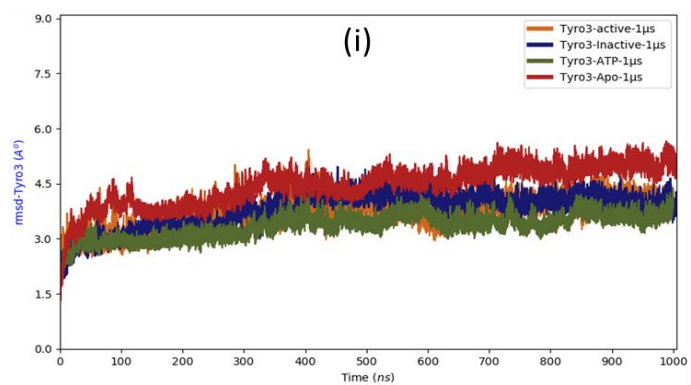
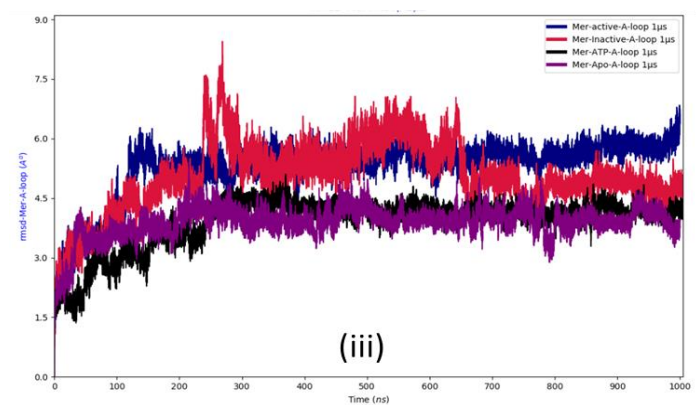
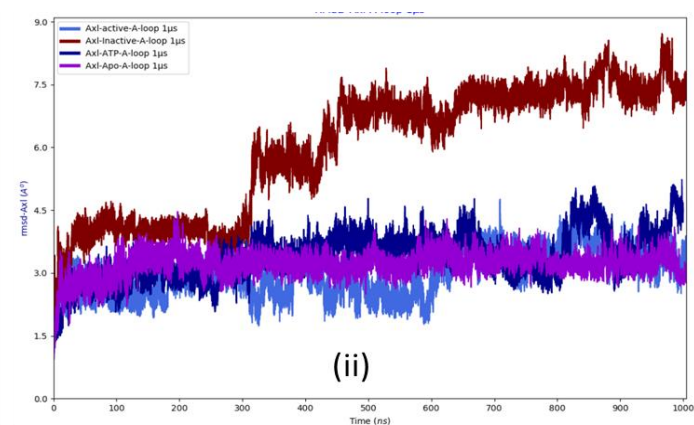
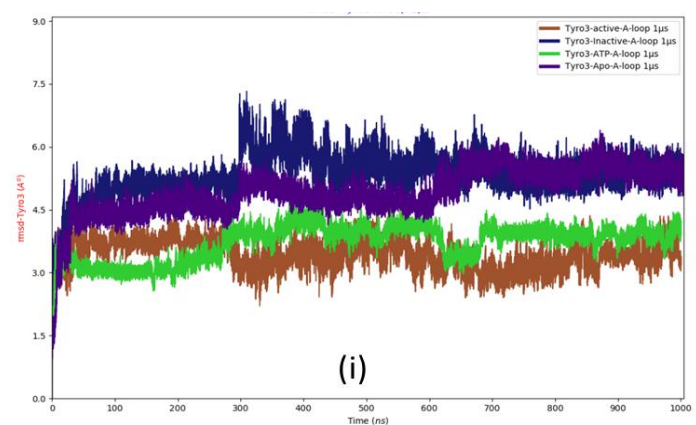


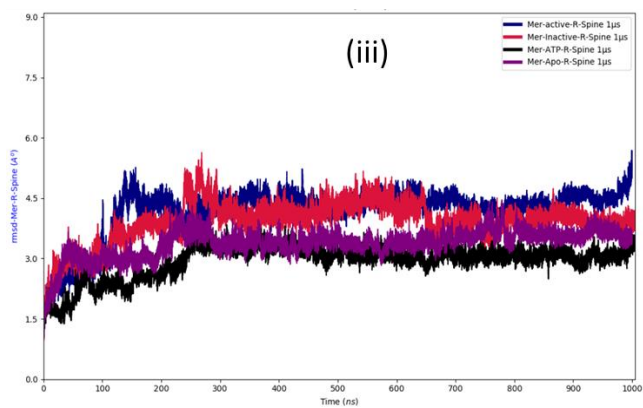
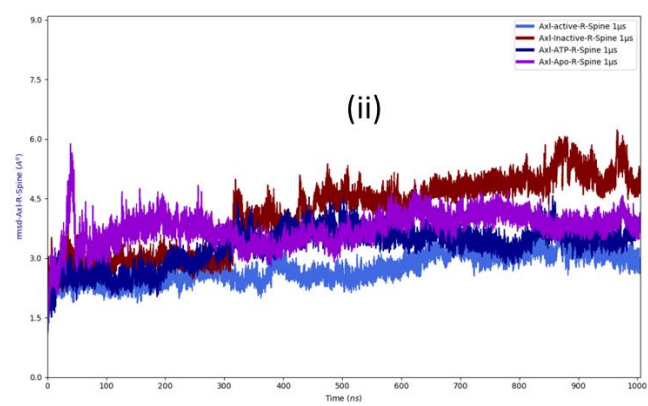
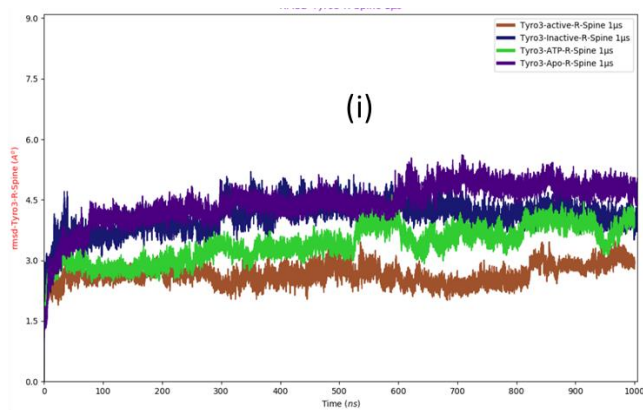
Figure 3.2) Regulatory spine analysis of cabozantinib bound active and inactive TAM kinase domains, apo and ATP bound active TAM RTKs at 1 μ s MD simulations. 3.2A) Tyro3; 3.2B) Axl; 3.2C) Mer



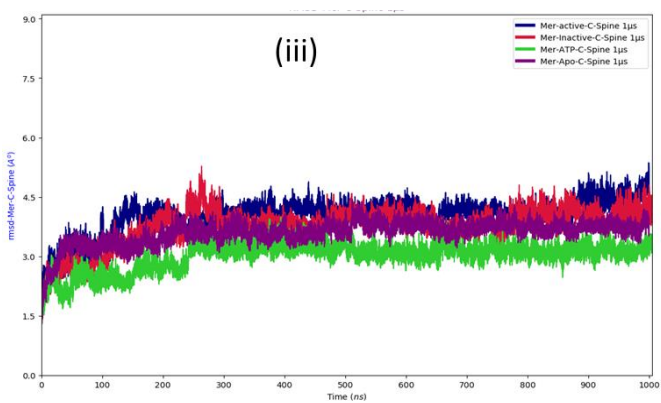
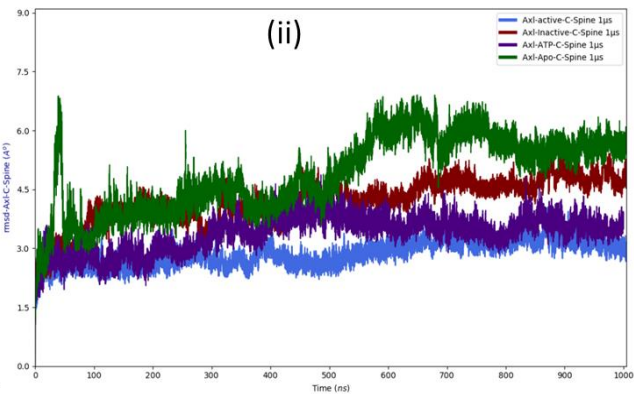
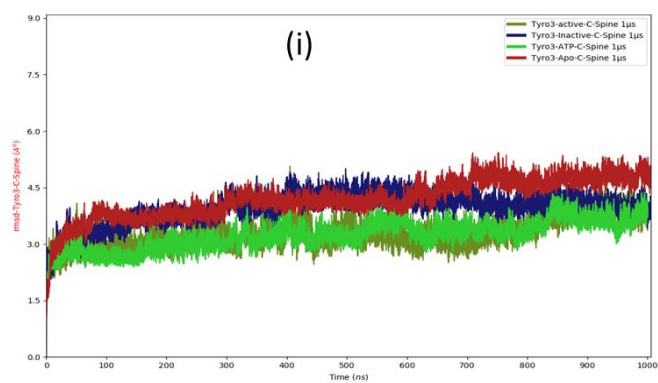
3A



3B



3C



3D

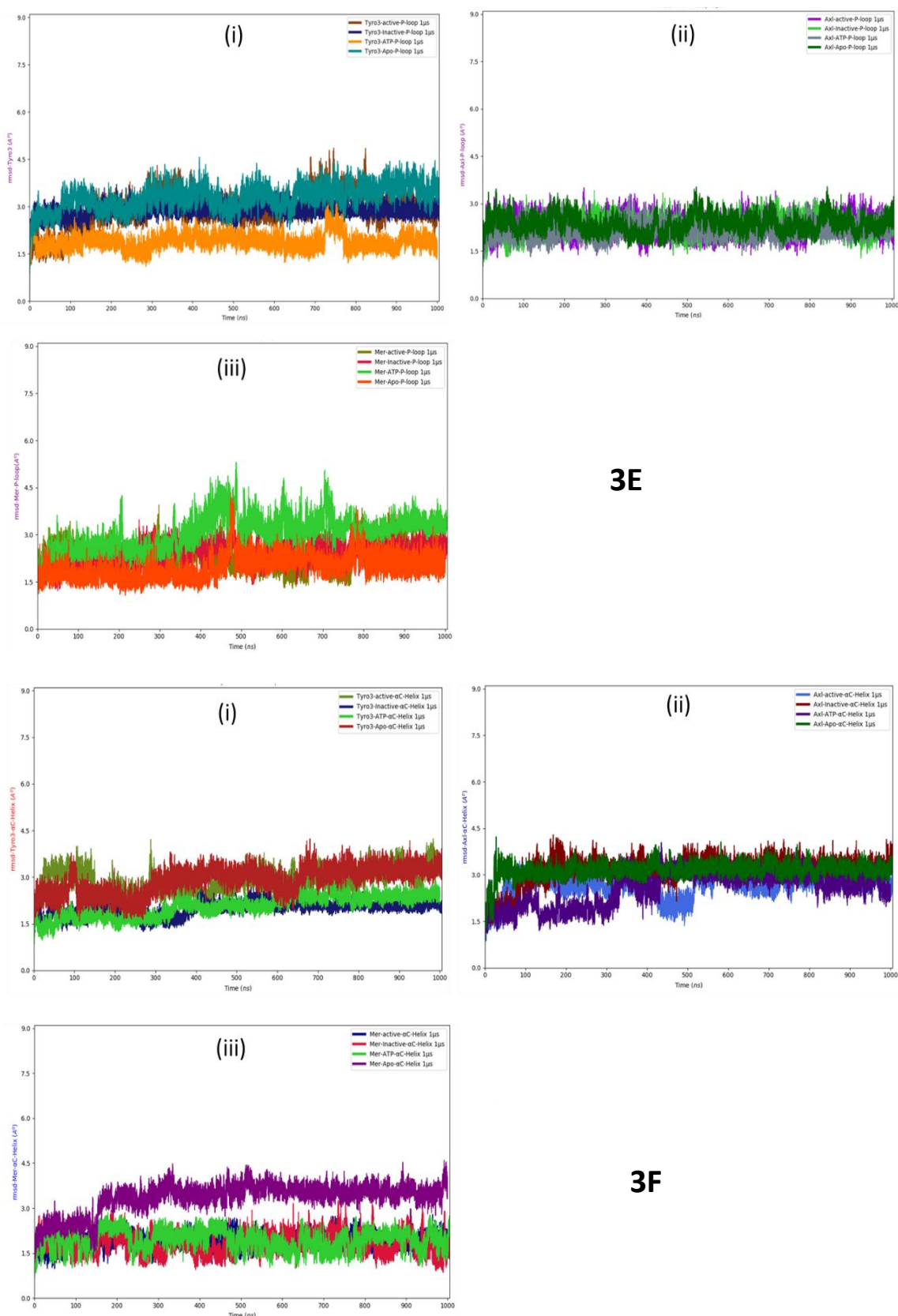
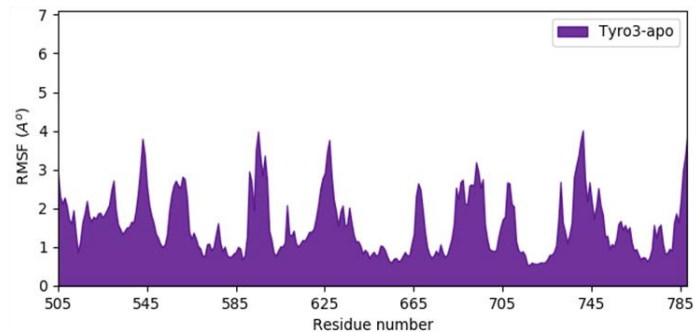
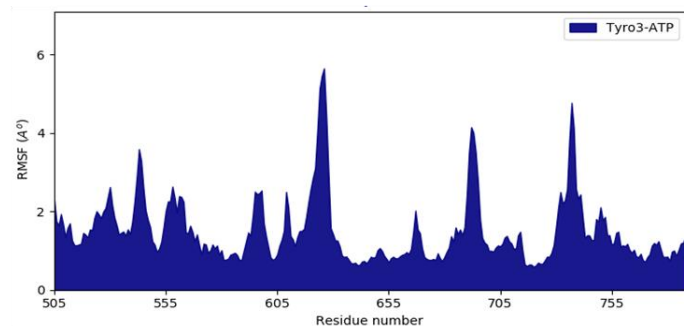
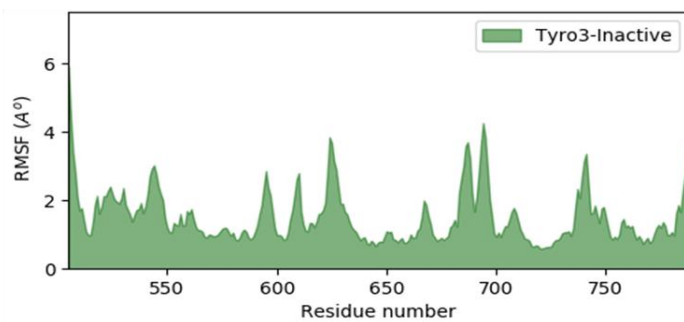
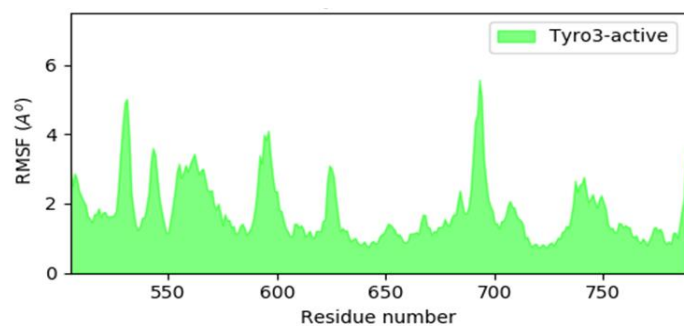
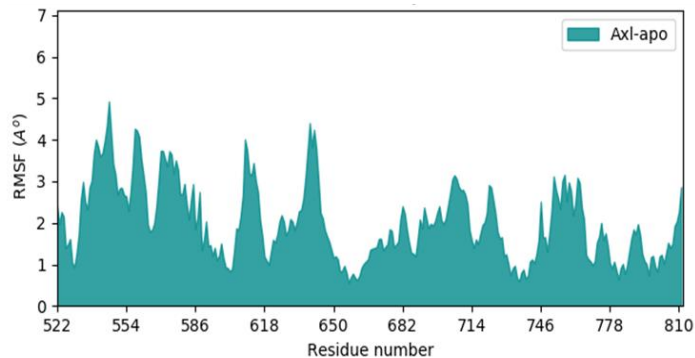
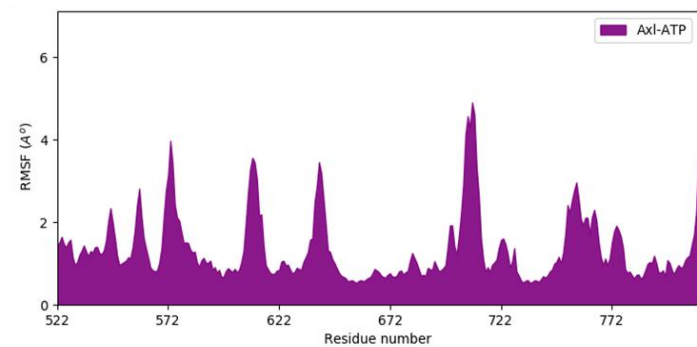
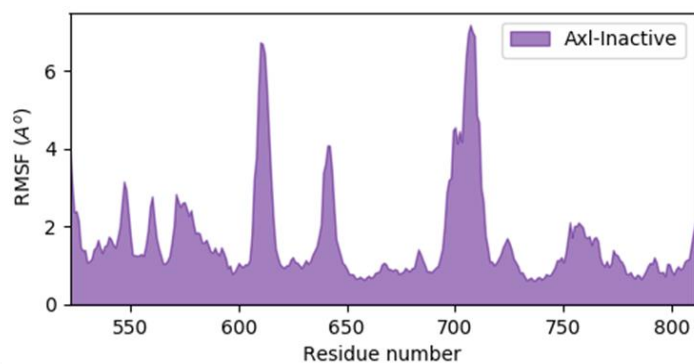
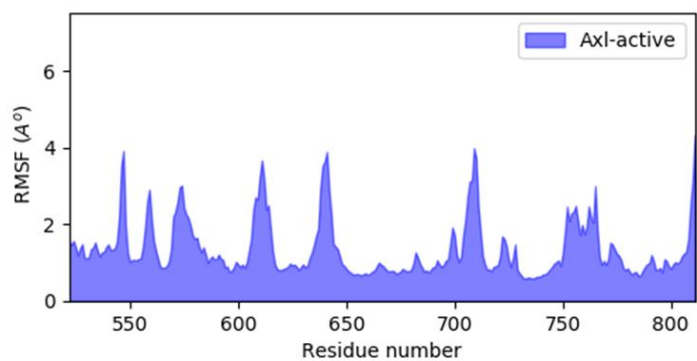


Figure 3.3) RMSD plots of apo, ATP bound TAM RTKs and cabozantinib bound active and inactive TAM RTKs from 1 μ s MD simulations. (3A) Tyro3, Axl, Mer C α atoms in the kinase domain (3B) Activation loop (3C) Regulatory spine (3D) Catalytic spine (3E) P-loop (3F) α C-helix

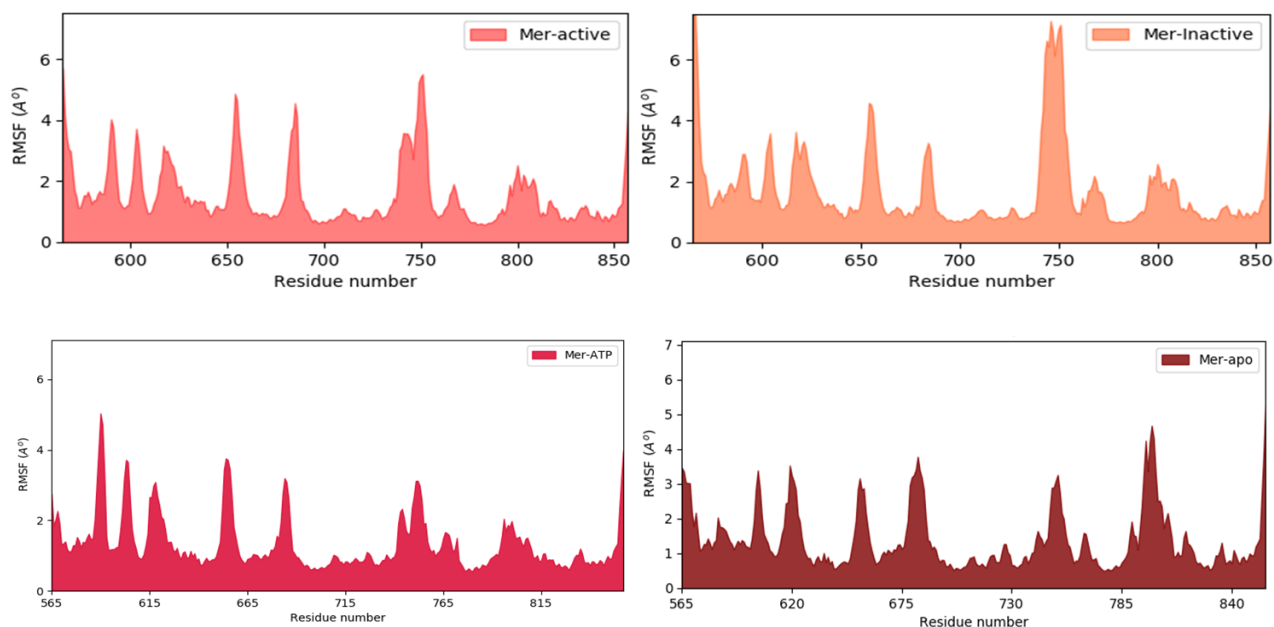
The RMSD plots are shown in Figure 3.3 A, B, C, D, E, F. The RMSD plots of protein C α atoms (Figure 3.3A), indicate that the structures converged at about 100 ns of MD simulations and the RMSD values lie within a narrow range from 2-4.5 Å. The ATP bound TAM kinase domains have lowest RMSD among all the systems studied. The TAM active state kinases form stable complexes when bound to cabozantinib. The apo Tyro3 and Axl have higher RMSD values among all systems studied. The RMSD analysis of specified regions in kinases are the key components to describe the distribution among inactive and active states. The regulatory spine of Tyro3 and Axl have well differentiated active states based on the lower RMSD (~ 2.8 Å) while the Mer active state has higher RMSD (~ 4.5 Å). All the inactive states of TAM kinases have an RMSD of 4 Å in the regulatory spine. The catalytic spine RMSD is higher in the case of apo Axl kinase but the ATP bound Mer and Tyro3 have lower catalytic spine RMSD (~ 2.8 Å). The RMSD of N-terminal P-loop are nearly similar in all the molecular systems studied. The RMSD of α C-helix region is distinguished among all TAM kinases studied and lie within a range of 1.5-3.0 Å. The active and apo states of Tyro3, the inactive and apo states of Axl, and the apo Mer kinases have higher and nearly similar RMSD values of the α C-helix among all the kinase states. The RMSD is lowest in the inactive Tyro3, Axl active, active and inactive Mer complexes. The RMSD of the activation loop is quite opposite to the α C-helix region. The active state Tyro3, active and apo states of Axl, and apo state of Mer kinase have lower and nearly similar RMSD values among all the systems. The activation loop in the apo and inactive Tyro3, inactive Axl, active and inactive Mer has highly dynamical conformation as can be seen from the higher RMSD values. Among all the systems studied, the inactive Axl activation loop is highly variable. The RMSD of regulatory and catalytic spines in the ATP complexes of TAM kinases is lower than 3 Å. The P-loop and α C-helix have lower RMSD (1.5 Å) and the RMSD of the activation loop is in between 1.5-3 Å. The cofactor ATP stabilizes TAM kinases with the adenine group coordinated at the hinge region of the kinase domain. The results from the RMSD are in correspondence with the RMSF plots (Figure 3.4 A, B, C). It can be seen that the hinge region is most stable in the Mer kinase domain when complexed with cabozantinib. From the analyses of the RMSD and RMSF plots, it is observed that TAM kinase domains have unique hidden dynamic states that can be distinguished from further analyses of MD trajectories.



4A



4B



4C

Figure 3.4) RMSF plots of TAM kinase domain from 1 μ s MD simulations. Kinase domains are numbered as per their primary structure. Axl indexing 539-553 (β 1- β 3 turn in the N-terminal domain-P-loop); 579-591 (α C-helix in the C-terminal domain); 689-724 (activation loop). 4A) Tyro3; 4B) Axl; 4C) Mer.

3.3.2. TAM kinase-cofactor complex activation pathway

The kinetic states appear due to the stereo-spatial arrangement of certain residues in specific α -helices, β -sheets and loop regions in the kinase domain. These kinetic states provide key insights into the activation of protein kinases in the presence of ATP and inhibitor bound to the active site. The apo, ATP bound and active/inactive Axl-cabozantinib molecular systems consist of well-defined kinetic state models during the MD simulations. The precise representation of the local spatial pattern in the active and inactive states of a kinase domain can be accessed via the regulatory spine and catalytic spine. The regulatory spine controls substrate molecule in the active site (α C-helix and activation loop). The catalytic spine regulates catalysis by allowing the ATP binding site at hinge region. The inactive kinase state should be converted into active state with the help of substrate binding at activation loop through the influence of regulatory spine hydrophobic residues which connect the dynamical movement of catalytic loop in α F-helix. The coordination between regulatory spine and catalytic spine evolve a dynamical conformation for the transfer of γ -phosphate from ATP to the substrate protein (Myers et al., 2019, Robinson, 2013, Kornev, 2006, Mohanty et al., 2016, Kim et al., 2017). The regulatory spine is continuous and linear in the case of normal metabolic kinase activity. The hydrophobic surface in the regulatory spine is vertically aligned (Leu-Met-Phe-His) in the apo form of all TAM kinases as can be

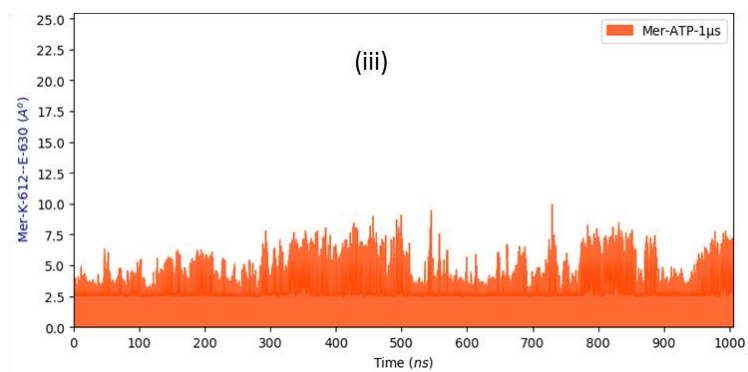
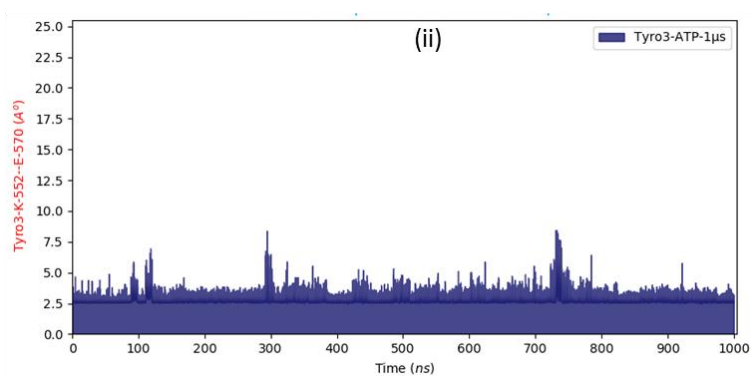
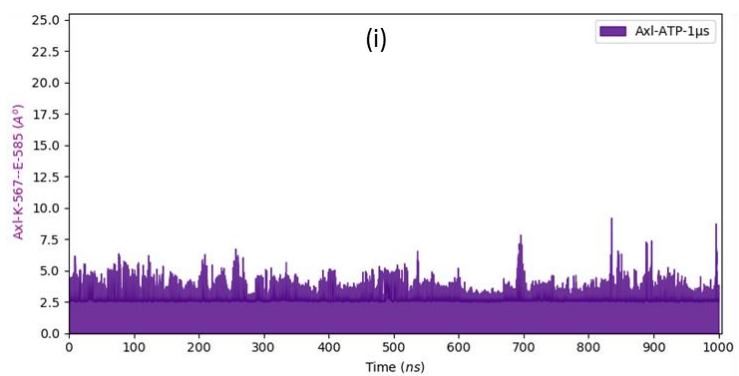
seen from the Figures (3.2 A, B, C). In the Axl active state, the regulatory spine is broken in the case of inhibitor bound form due to the expansion of activation loop that results in the extended space between α C-helix- Met589 and DFG motif -Phe691 at ATP binding site. The inactive Axl bound to cabozantinib has an intact regulatory spine due to the expansion of space between α C-helix Met589 and β_4 -strand Leu600 as a result of the outward rotation of α C-helix. The regulatory spine is retained in a similar way in the cabozantinib bound Tyro3 kinase domain in the active and inactive states. In the case of the cabozantinib bound active state Mer RTK, the regulatory spine fragmentation occurs between the P-loop, α C-helix and activation loop, whereas in the cabozantinib bound inactive Mer, the regulatory spine is retained. In the active site in Axl and Mer kinases, inhibitor occupies the shallow depth in between the α C-helix – activation loop, resulting in the broken regulatory spine. In the case of inactive states, the inhibitor binds at the hinge region of TAM kinases and therefore retaining the regulatory spine.

In the kinase active state, the regulatory spine is broken in Axl and Mer RTKs, whereas the catalytic spine is retained in the active state of Axl and Mer RTKs with no breakage in the hinge region. The regulatory and catalytic spines are coordinated in such a way that if the regulatory spine is broken, the catalytic spine is retained and vice-versa. The active Axl and Mer RTKs have broken regulatory spine but the catalytic spine is intact, but in the rest of the molecular systems the regulatory spine is intact and the catalytic spine is broken. It is like a lever pulling mechanism in the presence of higher concentration of inhibitor bound at regulatory site of kinase. The spine coordinated mechanism is important to ensure that the kinase is regulated from inactive state to active state mode in the presence of higher concentrations of substrate or high concentration of ATP in the cytosolic region. Therefore, the cabozantinib binding in the active state kinase influences at specified locations of the regulatory spine residues rather than catalytic spine. This can lead the catalytic spine to initiate catalytic activity towards passive mechanism to alert the body immune system with the help of chemokines. Whereas, in the inactive kinase state, the inhibitor binding to the regulatory active site or hinge region, regulatory spine activates either the dynamical movement of catalytic loop or catalytic spine to initiate the catalysis process with help of cofactor ATP. As a consequence, both the spines are well coordinated in the case of inhibitor bound to both active and inactive states to trigger apoptosis in malignant cells.

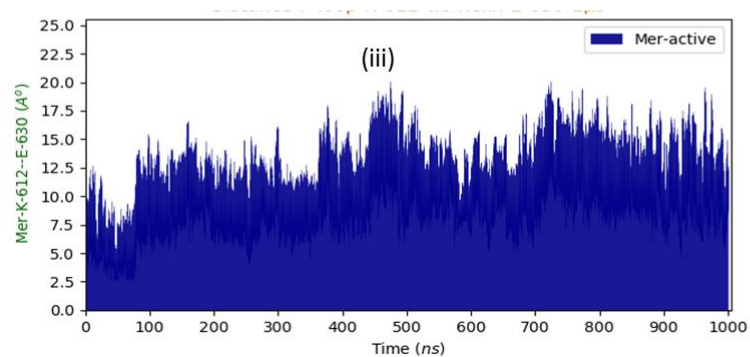
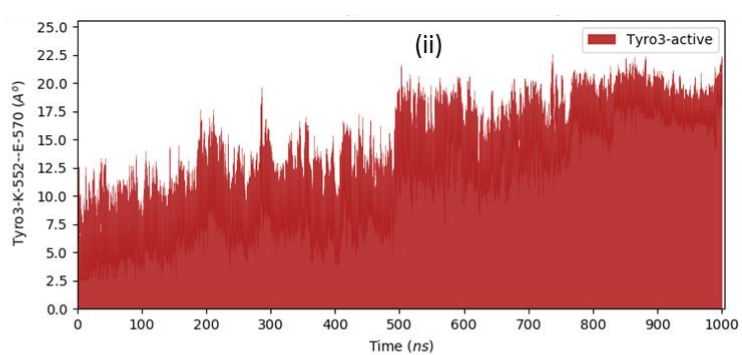
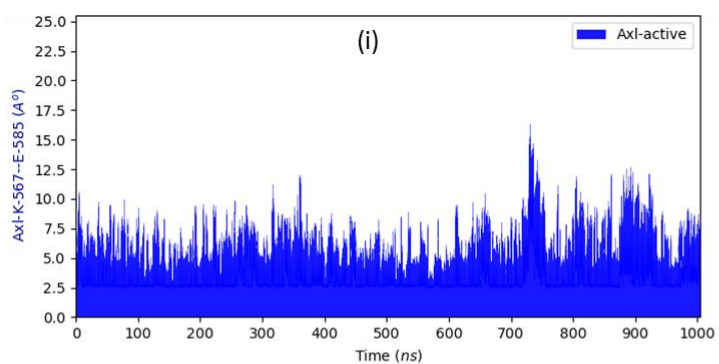
3.3.3. Confirmation of the existence of active states in ATP and active/inactive states in cabozantinib bound TAM kinases

A kinase domain can switch from active to inactive states and vice-versa due to either inhibitor binding, or influence of the regulatory spine and catalytic spine during MD simulations at longer timescales. Noticeable changes were observed in the spatial conformational states with inhibitor binding at the active site of the TAM RTKs. However, the specified regions of spatial orientations are not directly observable from the conventional RMSD plots. The inhibitor bound Axl kinase activation takes place in the transition

from active to inactive kinetic models. Therefore, these states coexist with broken regulatory spine in the active and inactive metastable states at the specified timescales of MD simulations. In addition, the RMSD of Axl differentiates due to the coexistence of active-inactive states throughout 1 μ s timescales (Figure 3.3A). The RMSD of specific loops in Axl is observed at higher square fluctuations occurring at the loop connecting β 4- β 5 strands (Glu609-Pro614), α C-helix and activation loop. It is evident from the RMSF plots that cabozantinib drug binding influences the inactive state of Axl and Mer kinases more than their active states (Figure 3.4). When cabozantinib binds the kinetically metastable states of TAM RTKs, it arrests the mechanism of kinase activity by inhibiting the up-regulation of its enzymatic activity. Regulatory spine is broken in inhibitor bound active state of Axl and Mer but it is intact in apo and cofactor (ATP) bound kinases. The active kinetic states of TAM kinase bound ATP at hinge regions shows Lys-Glu salt bridge distance retained within 4.0 $\text{\AA}$ range throughout 1 μ s MD simulations. This indicates that the ATP bound active TAM kinases retain their active state throughout the MD simulations. While only the cabozantinib bound active state of Axl has the salt bridge distance between P-loop and α C-helix, the Mer and Tyro3 kinases have longer distances (>7.5 $\text{\AA}$) due to the core expansion of activation loop region. In the case of inactive states of Axl, Mer and Tyro3 kinases these distances drastically increase beyond 12.5 $\text{\AA}$. This signifies that the Axl and Mer RTK kinetic models have well distinguished proportions of active and inactive states, while Tyro3 has similar ratios of active and inactive intermediate states (average 12.5 $\text{\AA}$ salt bridge distance between (P-loop) Lys and Glu (α C-helix) in Tyro3 active and inactive). The salt bridge distance between α C-helix Asp/Glu and Lys in activation loop of ATP bound states in Axl and Mer (< 5.0 $\text{\AA}$) and Tyro3 (>5.0 $\text{\AA}$) indicates highly dynamical structure than among all active and inactive states. These salt bridge distance analyses clearly differentiate cofactor (ATP) and inhibitor (cabozantinib) bound kinase domains at active site and active/inactive states, respectively (Figure 3.5 A, B, C, D). The inhibitor bound active/inactive kinase states are highly dynamical in nature than cofactor bound kinase states, therefore the inhibitor bound kinases might trigger apoptotic signaling pathways leading to inhibition. Based upon individual RMSD plots of the regulatory spine and activation loop (Figure 3.3B, 3.3C), it can be seen that cabozantinib binding influences the activation loop and hydrophobic spine in individual kinetic states. A specific spatial conformational variation in RTKs occurs only in the activation loop and regulatory spine. The active and inactive forms of apo and active ATP bound conformers of TAM kinases appeared to have intact regulatory spine. This could lead to the normal signal transduction process while the regular ligands [GAS-6 and Pros1] bind to the extracellular regions of TAM RTKs.



5A



5B

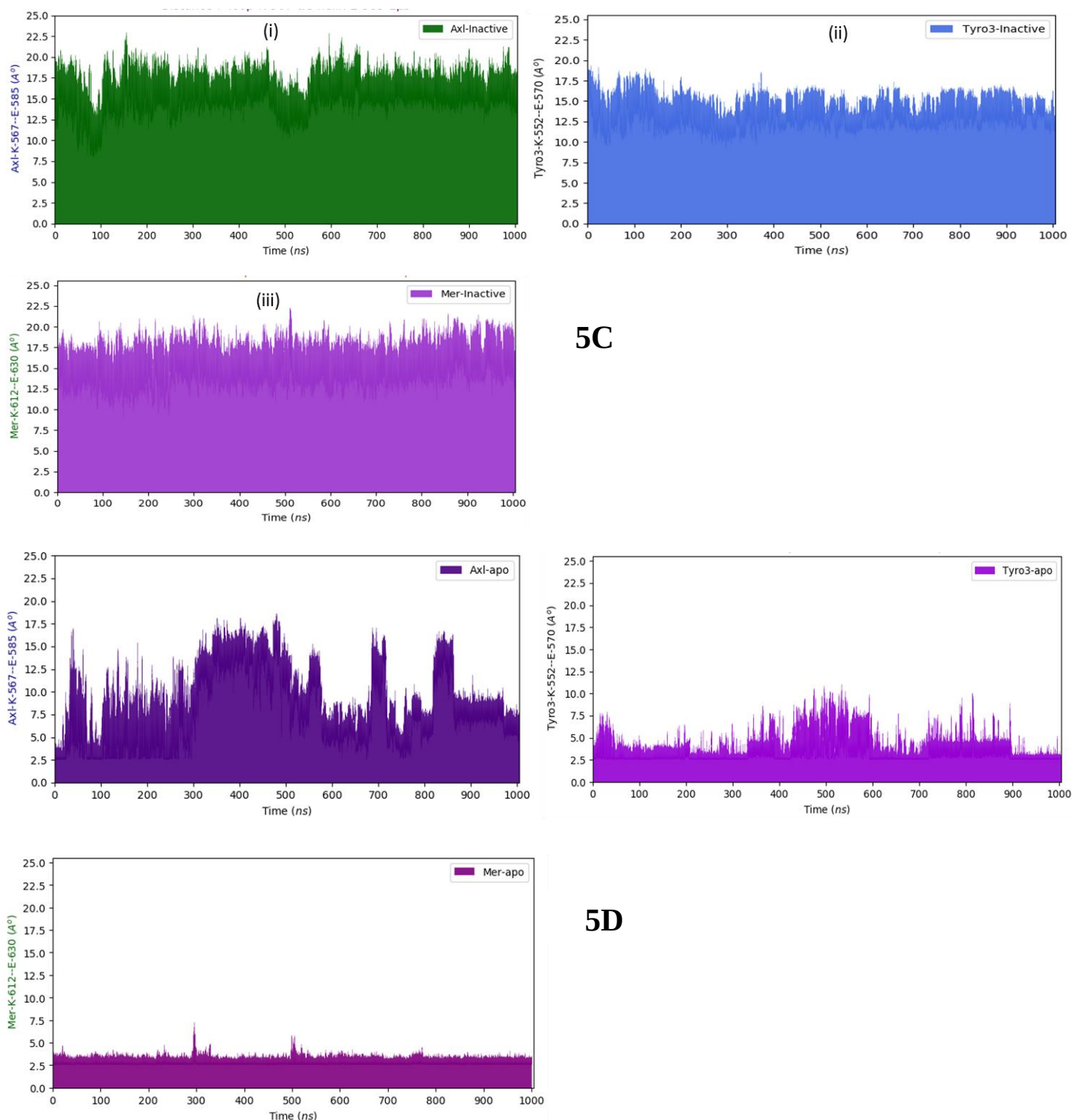


Figure 3.5) Distance plots between side chains of Lys P-loop - Asp α C-helix pairs in apo, ATP and TAM-cabozantinib bound active and inactive kinase domain from 1 μ s MD simulations. Axl (K567 – E585); Tyro3 (K552 – E570); Mer (K612 – E630); (A) ATP bound; (B) Active inhibitor bound; (C) Inactive inhibitor bound; (D) Apo state

The analysis of salt bridge distance between the activation loop and α C-helix reveals the hidden conformers among the apo, ATP and inhibitor bound TAM kinases. The salt bridge interaction in the apo kinase is retained within a reasonable distance between the α C-helix and activation loop residues Asp581-Lys695 (3.99 Å, Axl) or Glu626- Lys739 (3.86 Å, Mer) or Glu566-Arg680 (3.44 Å, Tyro3). The salt bridge distance between Asp/Glu (α C-helix) - Lys/Arg (activation loop) in cabozantinib bound active states increases in Axl and Mer RTKs due to the expanded core in the inhibitor binding site in RTKs (Tyro3-3.16 Å; Axl-13.92 Å; Mer-10.28 Å), but in the inactive states of TAM RTKs salt bridge distance between α C-helix and activation loop is lower for Axl RTK (Tyro3-6.82 Å; Axl-2.83 Å; Mer-8.83 Å). These salt bridge distances provide support to the stationary state distribution in apo TAM RTKs. These salt bridge distances in the ATP bound TAM kinases is observed to be greater than 5 Å (Tyro3-7 Å; Axl-10.0 Å; Mer-12.0 Å).

The salt bridge is retained in the apo form, active states of Tyro3 and inactive states of Axl. The salt bridge distance analysis provides a clear evidence that the kinases coexist in active and inactive state models while binding with inhibitor at the active site. The large distance across the regulatory site of kinase active states occurred due to a β -sheet formation in the activation loop and inward rotation of α C-helix. This causes the extended nature of regulatory active site between α C-helix and activation loop. The inactive state models have α C-helix outward rotation and activation loop undergoes shift to helical structure to minimize the active space across α C-helix and C-lobe in the RTKs. These results provide further support to regulatory spine analysis. But most of the active states in Axl and Mer forms have broken regulatory spine between α C-helix and activation loop therefore the distance between these domains is extended and the salt bridge interaction is disturbed due to the increased distances between Asp/Glu (α C-helix) - Lys (activation loop). In the inactive state of TAM RTKs, the regulatory spine is reinstated due to the bound inhibitor at the hinge region of kinase and expansion of space between P-loop and α C-helix. In overview, in the cabozantinib bound TAM kinases, the salt bridge distance is higher in active states than inactive states, as the distance of salt bridges Asp/Glu (α C-helix) - Lys/Arg (activation loop) in the active states are above 10 Å, the inactive states have below 10 Å.

3.3.4. Post-MD data analysis of TAM RTK kinase domain

The preliminary MD simulations data acquired from AMBER trajectories were analyzed to ensure that kinetically active and inactive states were investigated with the help of PCA. PCA analysis was carried out on 1K conformer samples of trajectories out of 40K for clear visualization of data points from kinetic transition states in the active to inactive kinases. The histogram showed that the random distribution of kinase state trajectories data was extrapolated as training and test sets of individual components validated

with shuffle-split cross-validation in PCA plot. All apo and inhibitor bound forms of TAM RTKs have random distribution of states that are very unique in nature from the respective scatter plots of kinase trajectory analysis. This is a preliminary analysis to propose the hidden dynamic states existing in longer timescale MD simulations and trajectory data of kinases.

The metastable kinetic models were built based upon advanced trajectory data analysis using python based scripts. All TAM trajectories data was sampled into vectorized and clustering was done using Keras-state algorithm for MSM model generation (Prinz et al., 2011, Schwantes and Pande, 2013, Harrigan et al., 2017). The MSM data of TAM kinases was bootstrapped from 1 μ s of trajectory data to generate HMMs to reveal the unfolding and refolding of the activation loop from active state to inactive states. The metastable trajectories are well converged as shown by VAMP score. Discrete clustering of protein backbone state distribution featurisation was performed to show distinct kinetic stable states in all TAM kinases. All HMM states are key intermediate conformers to describe the kinase inhibitory activity when bound to cabozantinib. As per the analysis of metastable kinetic state forms, higher numbers of active state models are present in Axl and Mer than the number of kinetic transitions states of inactive forms. However, the Tyro3 has approximately similar numbers of kinetic state models in their respective active and inactive states which are included in state distribution plots. The MFPT error bars were validated with Bayesian HMM model validation with lag time of 50 states. From these analyses it is inferred that Tyro3 RTK kinase domain states have combined and coexisted metastable state transitions among the active and inactive forms rather than the dominance of either the active or inactive kinetic states as observed in Mer and Axl RTKs. Therefore, the Tyro3 has more intermediate states than Axl and Mer. The MFPT values of Tyro3 indicate that activation and deactivation occur in equal ratio below 100 ns; whereas the Axl and Mer have different activation timescales (after 200 ns) and their deactivation takes place around 100 ns timescales; The influence of these major changes in the kinase domain is due to the conversion of active to inactive states through kinetic transition metastable equilibrium states.

In the inhibitor bound form of TAM kinases, greater state distribution models coexist in the active forms than in the inactive forms. The drug bound to kinase active state influences the kinetic signaling pathways more rather than the inactive state (Roskoski, 2015, Taylor and Kornev, 2011, Sultan et al., 2018, Shukla et al., 2014). Therefore, the active state kinase bound to inhibitor is more susceptible to arrest the dysregulated kinase activity (shown by the broken regulatory spine) in all kinetic HMM states. These observations provide key insights to describe that the kinase activity can be arrested through active state models of inhibitor bound RTK, where regulatory spine breaks in between activation loop and α C-helix in the active states (Robinson, 2013, Parsons and Parsons, 2004, Bowman and Pande, 2010). The hydrophobic surface regulatory spine is retained in the apo form of all the three TAM kinases. The regulatory spine is retained in the cabozantinib bound Tyro3 RTK in the active and inactive states, due

to the increased distance between P-loop and α C-helix. This retaining of regulatory spine in Tyro3 RTKs indirectly influences the number of active and inactive state distribution in equal proportions. In the inactive Axl and Mer RTKs, intact regulatory spine is observed due to increase in the distance between the P-loop and α C-helix, whereas in the active Axl and Mer RTKs, regulatory spine fragmentation occurs between α C-helix and activation loop, due to the lower distance between P-loop and α C-helix. These observations are shown in Figures 3.2 (A - C). The discrete clustering of MSM estimation and validation was done with reversible estimation equilibrium transition probabilities. The discrete kinetic state models were further validated by analysis of hidden markov kinetic models. The implied relaxation timescales are extracted to validate the HMM in order to ensure the conditional transition probabilities among 250 microstates. Therefore the implied timescale analyses indicated that the kinetic state distribution occurred within time intervals of a few nanoseconds range among 1 μ s MD simulations timescale.

The Mer active states have longer MD kinetic relaxation timescales among the active MSM kinetic forms of TAM RTKs. The inactive Axl kinetic state models have higher relaxation timescales within short range of time intervals. The critical observation from all TAM apo and inhibitor bound active and inactive kinetic states implied from timescale plots, with 4.5 ns timescale separation as the average implied relaxation timescale among all. The Tyro3 apo has more relaxation time intervals than the rest of kinase systems. The kinetic relaxation time intervals revealed that the inhibitor bound TAM RTKs showed kinetic metastable state transitions due to various periodic time laps even though all TAM RTKs are bound with same inhibitor (cabozantinib).

The MSMs of the members from same class of protein kinase complexes (TAM kinases bound to cabozantinib) is expressed as different relaxation timescale intervals obtained from the MD simulations. The free energy and stationary state distribution of apo Axl is higher than Tyro3 and Mer. From the Table 3.1, it is inferred that there are unique kinetic Markov state models existing among them. These are classified as “kinetic non-equilibrium transition state models” (Tyro3 apo, Mer active, Axl inactive). This is further discussed in kinetic transition analysis. The lowest free energy and equal stationary distribution exist in stable kinetic model states of TAM kinases (Axl-active, Mer-inactive). The kinetic transition states between Axl active and Mer inactive has higher free energy and approximately equal stationary distribution values (Tyro3 active/inactive) and are classified as “kinetic equilibrium transition state models”. As per the state distribution difference between active-inactive states of Axl inactive HMM has half (1/2) of the stationary distribution of Axl active (more active state distribution). The inactive Mer has $\frac{3}{4}$ of the state distribution of active Mer RTK. The Tyro3 has equal contribution in active and inactive stationary distributions among kinetic HMM states. The surface free energy of Axl has same energy values in the active and inactive states (~ 4.0 kcal/kT per 5 states-Axl) but Tyro3 and Mer have 0.5 kcal and 1.2 kcal, respectively per five MSM states energy difference between the active and inactive hidden

Markov states. Each hidden MSM state contains five metastable kinetic conformers from sampling of 40K conformers to study the MSM validation.

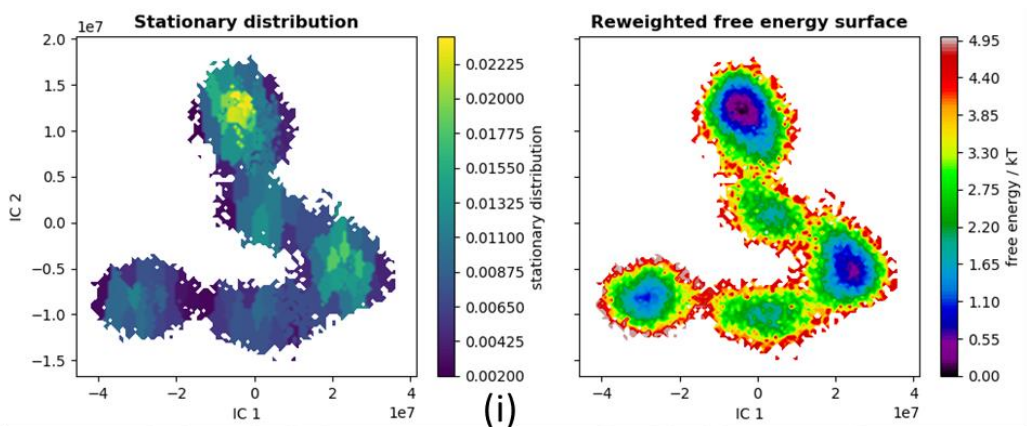
Kinetic metastable states	Kinases types		Axl		Mer		Tyro3	
		STATES	π	G/kT (kcal per HMM state)	π	G/kT (kcal per HMM state)	π	G/kT (kcal per HMM state)
	Active	1	0.080859	2.515052	0.072653	2.622066	0.000000	inf
		2	0.000000	inf	0.095574	2.347855	0.210675	1.557439
		3	0.328240	1.114010	0.271887	1.302367	0.094495	2.359208
		4	0.206730	1.576340	0.286168	1.251176	0.343637	1.068168
		5	0.384171	0.956668	0.273718	1.295657	0.351193	1.046420
Transition states				S4-S5		S2-S4		S4-S5
		STATES	π	G/kT	π	G/kT	π	G/kT
	Inactive	1	0.128210	2.054082	0.032751	3.418814	0.030226	3.499054
		2	0.262320	1.338189	0.037707	3.277907	0.058256	2.842902
		3	0.138915	1.973891	0.260523	1.345064	0.132997	2.017429
		4	0.171230	1.764746	0.151968	1.884087	0.255189	1.365750
		5	0.299324	1.206230	0.517051	0.659614	0.523331	0.647540
Transition states				S2-S5		S4-S5		S4-S5
		STATES	π	G/kT	π	G/kT	π	G/kT
	Apo	1	0.008222	4.800930	0.081258	2.510125	0.078556	2.543949
		2	0.052043	2.955685	0.239732	1.428232	0.083054	2.488261
		3	0.080056	2.525032	0.131020	2.032406	0.139106	1.972522
		4	0.165858	1.796625	0.228406	1.476629	0.293660	1.225334
		5	0.693821	0.365541	0.319583	1.140738	0.405625	0.902326
Transition states				S4-S5		S4-S5		S3-S5

Table 3.1) Tyro3, Axl and Mer kinetic transition states analysis with specified free energy of nine HMM states. Red color indicates metastable kinetic equilibrium transitions states. Blue color indicates metastable kinetic non-equilibrium transitions states

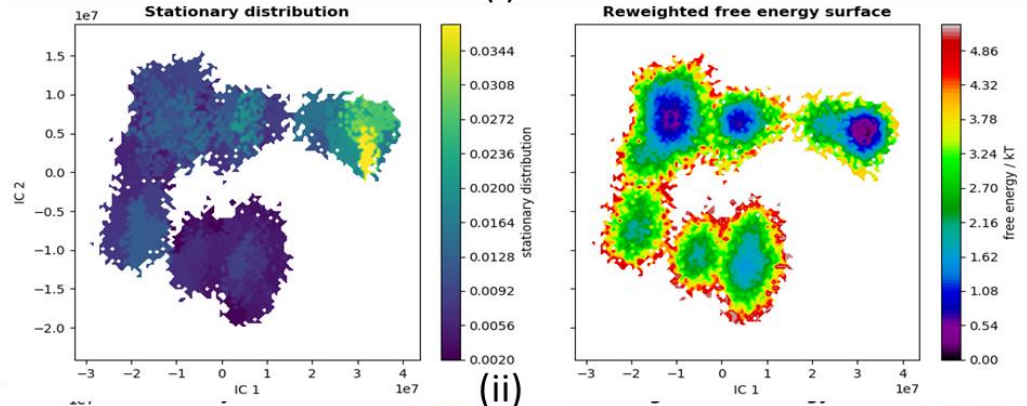
3.3.5. Kinetic transition state analysis

The estimated five state kinetic metastable models were designed based upon active space distribution of HMMs in TAM RTK kinase domains. All the five metastable state transitions occurred based upon kinetic transition energy (Weinan and Eijnden, 2010, Salvalaglio et al., 2014, Metzner et al., 2009) (Table-3.1). The apo Axl has higher transition energy (4.8 kcal), inactive Mer (3.4 kcal) and inactive Tyro3 (3.5 kcal). Out of the nine kinetic states, six kinetic transition states are represented as metastable kinetic equilibrium transition states as these kinetic transitions occurred in S_4 - S_5 states. The metastable kinetic non-equilibrium transitions exist in various types of kinetic metastable states (Tyro3 apo – S_3 - S_5 ; Mer-active- S_2 - S_5 ; Axl inactive- S_2 - S_4) from the nine metastable transition states (Figure 3.6 A, B, C). All non-equilibrium kinetic transitions occur with a very low transition energy (2-2.6 kcal). These hidden states are classified based upon kinetic transition energy and state transitions. All the metastable kinetic equilibrium transitions occurred with a high energy (2.3-4.8 kcal). As per the individual TAM RTK, the Axl apo kinase has higher kinetic transition energy among all TAM RTKs in the apo and inhibitor bound active and inactive forms. The next higher kinetic transition energy exists for Mer and Tyro3 inactive forms. It is evident that all inhibitor bound RTKs exhibit different kinetic metastable states in the overexpressed RTKs during the protein function. According to approximate difference in transition probability of active to inactive metastable kinetic states in Tyro3, Axl, and Mer RTKs, for Tyro3, 1st MSM state has higher transition probability difference (50 %), for Axl and Mer RTKs, 2nd MSM states have higher transition probability difference. The transition of kinase active state to inactive state can be explained based upon kinetic metastable states of these specified MSM conformer analysis (Husic and Pande 2018). From the Figure 3.7, it can be seen that the stationary state distributions in Axl active are doubled when compared to Tyro3 active, and Mer inactive states has only 4/3 proportion. Therefore, Axl active RTK has more active stationary states. The relative transition state probability is explained on the basis of salient feature analysis in hidden Markov kinetic states. These could be key intermediate structures among subfamily of TAM RTK kinase domains. However, these protein kinases driven from active to inactive states expressed significant structural changes upon binding with inhibitor. The active state of Axl kinase consists of activation loop that transits from β -sheet to α -helical structure in the inactive state (Figure 3.7A and B). Mer RTK shows high structural changes in the activation loop which converts from loop (active) to helical (inactive) in their respective state transitions (2-2 transition probability), while Tyro3 does not have any significant change in the MSM kinetic states (Figure 3.8A and 3.8B).

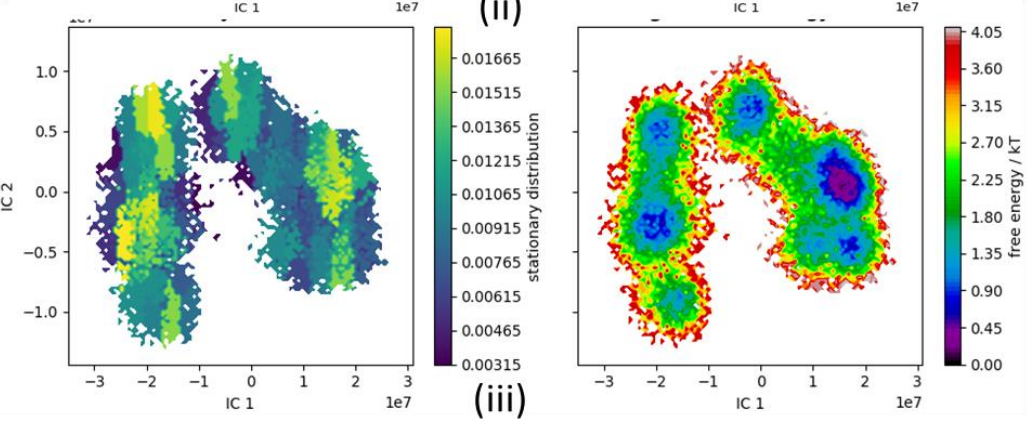
Tyro3-apo



Mer-active

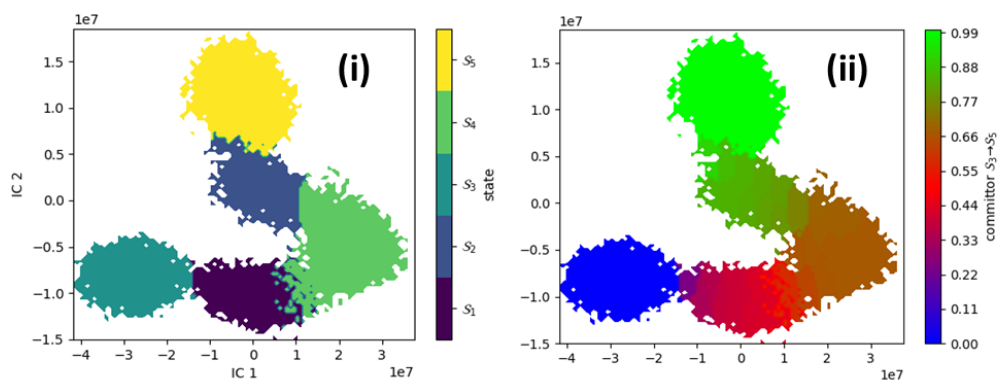


Axl-Inactive

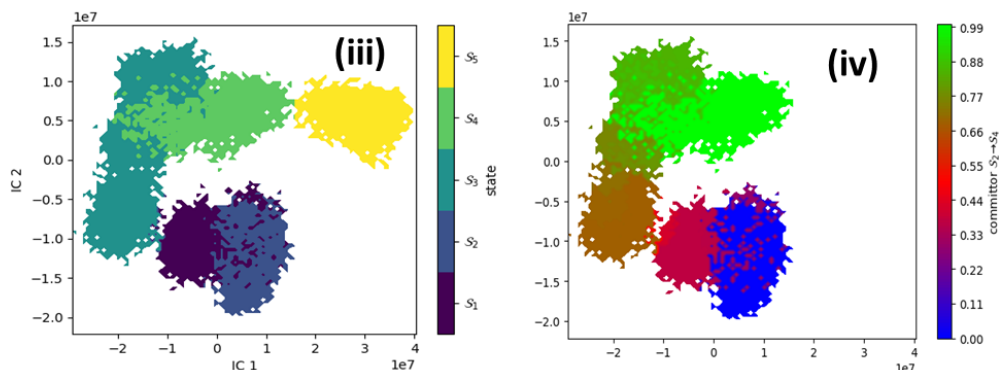


6A

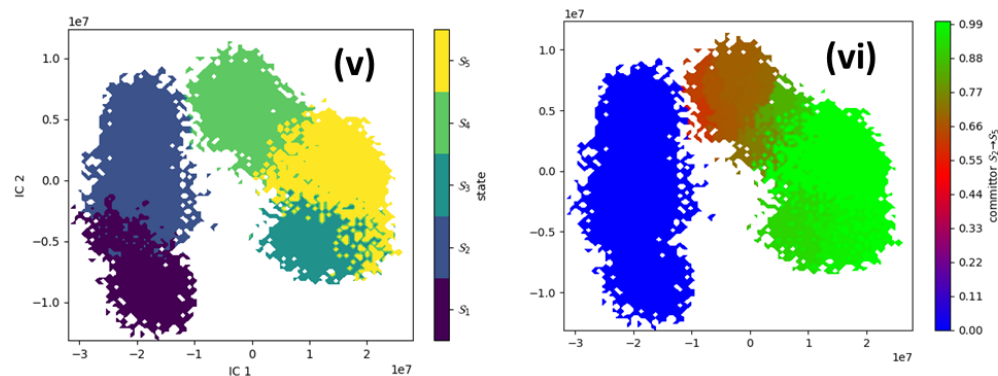
Tyro3-apo



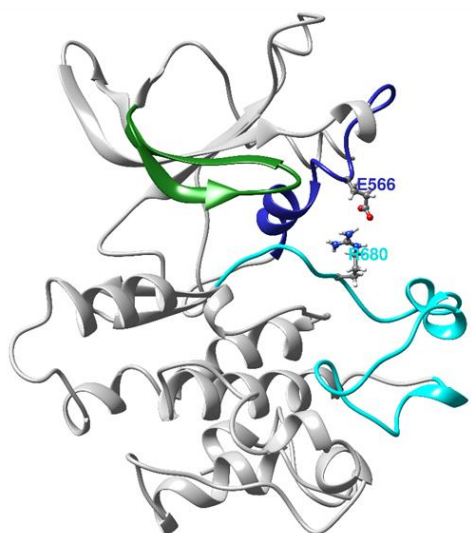
Mer-active



Axl-Inactive

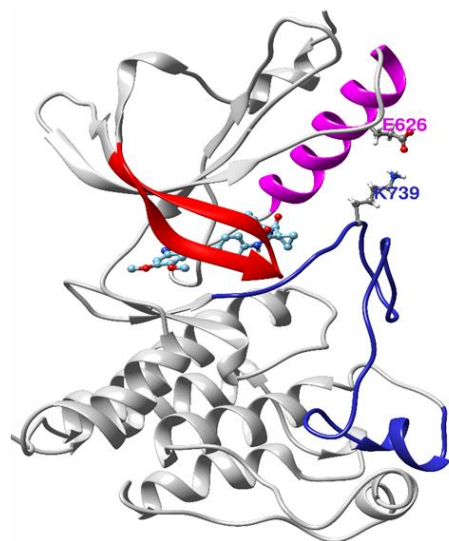


6B



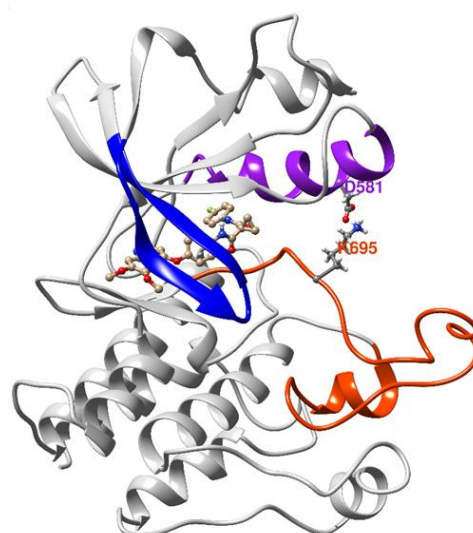
Tyro3 apo

I



Mer active

II



Axl Inactive

III

6C

Figure 3.6) Kinetic metastable five states estimation and kinetic metastable transition state analysis of metastable kinetic non-equilibrium transitions states Tyro3 apo and cabozantinib bound Mer-active and Axl-inactive states of RTK from 1 μ s MD simulations.

(3.6A) Stationary states and reweighed free surface energy of non-equilibrium transitions states

(3.6B) MSM five states estimation and kinetic transition states;

(3.6C) Specific states distance between side chains of Asp/Glu- α C-helix – Lys/Arg-activation loop pairs in Tyro3 apo and cabozantinib bound Mer-active and Axl-inactive; (I) Tyro3 apo; (II) Mer active; (III) Axl inactive.

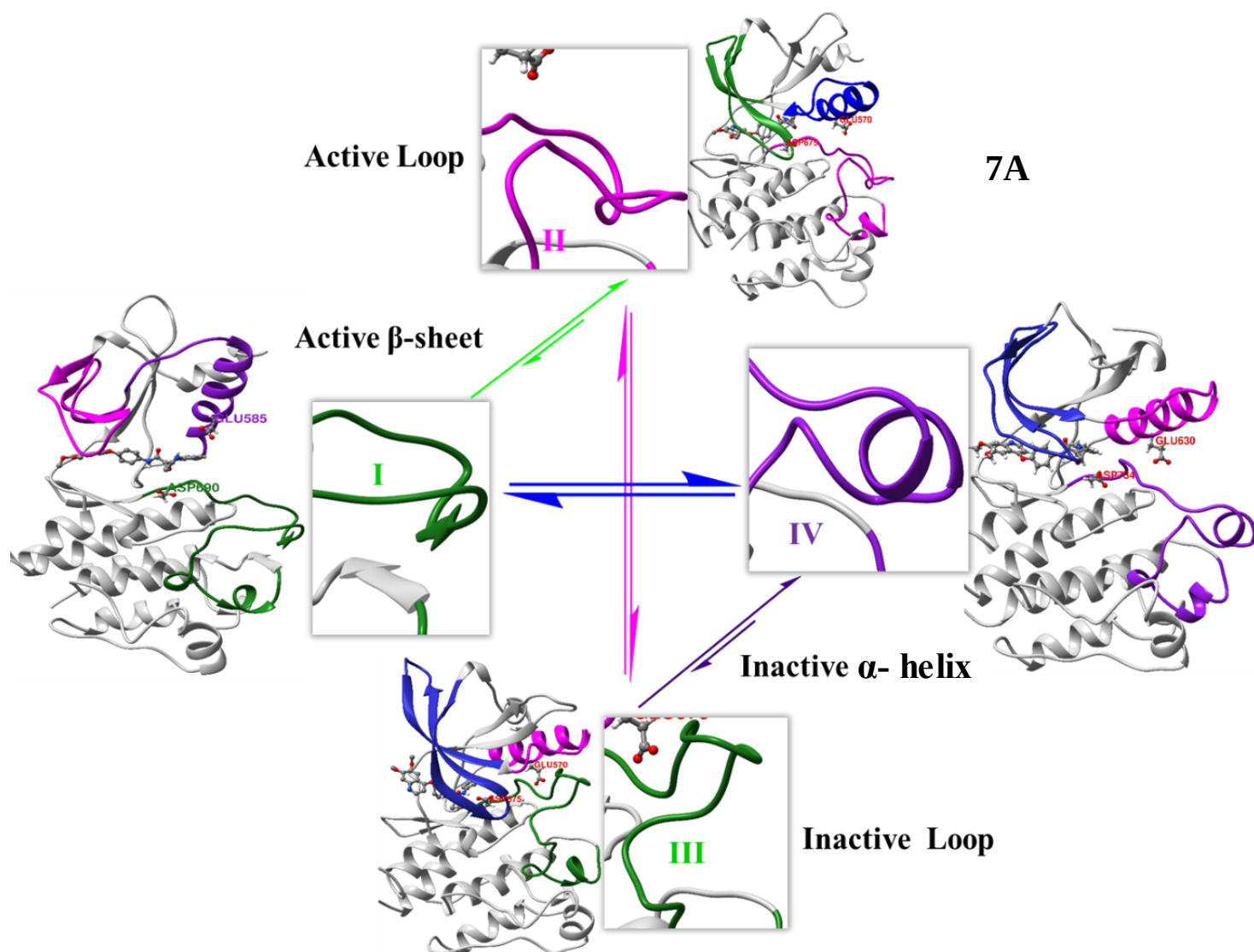


Figure 3.7A) Schematic view of three kinetic equilibrium metastable states among HMM states involved in allosteric activation and deactivation from active to inactive states in RTKs.

3.3.6. Mechanistic strategy of TAM RTKs activation when complexed with cabozantinib

The dynamical movement of the regulatory and catalytic spine residues are a result of the coordinated alterations in the kinase structural domain during the cellular signal transduction process. The selective kinase inhibitor (cabozantinib) arrests the activity of these overexpressed kinase domains via the dynamic movement of both these spines and distancing the space between α C-helix and activation loop. This can be supported from the results of distance plots shown in the active states of Tyro3, Axl and Mer that have undergone large expansion of protein core between α C-helix- activation loop in regulatory active site. Therefore, the kinase activation is carried out by the active state modes. The five metastable states from Chapman-Kolmogorov test described transition probability from 40K frames of dynamic kinetic metastable states for each of the protein complex trajectories, was obtained from AMBER MD data with 95% confidence level. Combining all these transitions probabilities with transition states and assigning five sampled metastable states could provide good insights and predict long lived transition states in the MD simulations trajectories with Perron-cluster cluster analysis (PCCA++) clustering algorithm (Scherer et al., 2015, Schwantes and Pande, 2013, Harrigan et al., 2017, Perez-Hernandez et al., 2013). The RTKs are involved in signal transduction process in which dysregulated kinase is inhibited such that the cells initiate programmed cell death with the help other proteases belonging to the caspase enzyme (Kim, et al., 2017). The regulated and dysregulated kinases can be distinguished with help of regulatory spine (Leu-600 { β 4-sheet P-loop}; Met-589 { α C-helix}; Phe-691 {DFG- activation loop}; His-670 {catalytic loop}) closed and open conformers of apo, ATP and inhibitor bound TAM RTKs, respectively, due to significant conformational changes. A regulated kinase has closed and continuous regulatory spine in both active apo and ATP bound form in RTKs. The hydrophobic surface is in a closed manner and continuous in apo and ATP bound form of active Tyro3, Axl and Mer RTKs. This space has expanded in the case of active states of normal physiological kinase mechanism. This is achieved based upon activation loop refolded into β -sheet and π -stacking with β -sheet structure of the catalytic loop in active state model (observed in Axl). But inhibitor bound dysregulated kinase experiences large conformational deformations in their regular structures due to the influence in certain parts of RTKs with overwhelmed hidden dynamic states to trigger kinase domain equilibration between active and inactive states. Indeed, the drug (cabozantinib) bound at RTKs active site, triggers the activation loop folding into either β -sheet (Axl active-state) or α C-helix (Mer inactive-state) (Figure 3.7A). These kinetic metastable states have transition from active to inactive states through intermediate structure (transition-state) and vice-versa (Weinan and Vanden-Eijnden, 2010, Metzner et al., 2009, Noe and Clementi, 2015). The dynamic states would proceed through mechanistic pathways to initiate signaling process as expanding or compressing

of the activation loop, outward/inward rotation of α C-helix and extended movements in the P-loop. It can be seen that the active site cavity is enhanced in the presence of inhibitor bound active state that has broken regulatory spine obtained by moving apart the Glu residue on α C-helix and Phe residue in DFG motif associated with the activation loop. The uncertainty of migrated residues could be withheld in a particular state of kinase domain vertically from N-lobe towards C-lobe. The regulatory spine is broken in the active state only in situ with all four residues moving away from the broken hydrophobic surface between α C-helix bound Met-589 and DFG motif bound Phe-691 due to the extend space of activation loop and inward rotation of α C-helix. Inactive state model kinase consists β 4-strand bound Leu-600 and α C-helix bound Met-589 in situ regulatory spine intact in a continuous manner due to the outward rotation of α C-helix and the activation loop has recoiled into α C-helix where DFG motif and α C-helix moves away from the P-loop of β -sheet (Figure 3.7B).

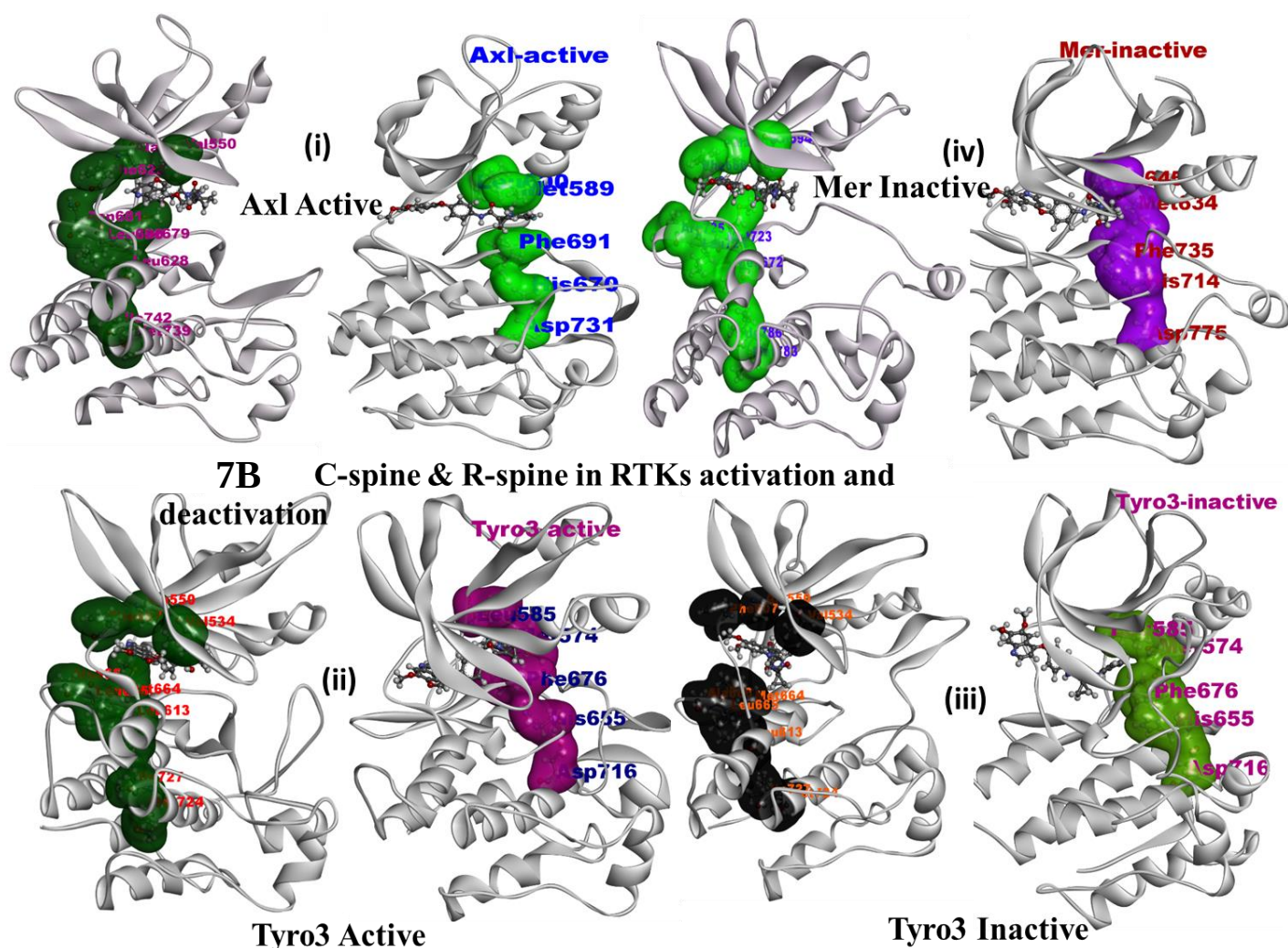
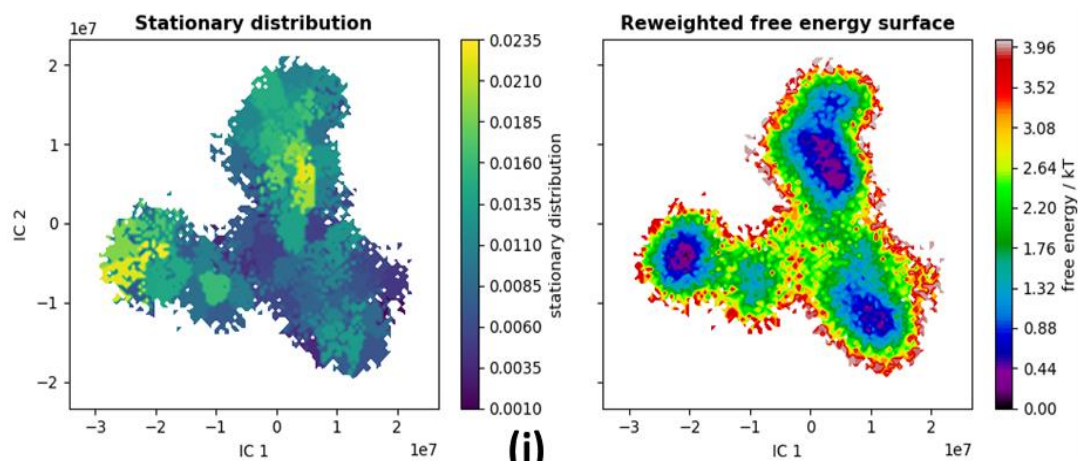


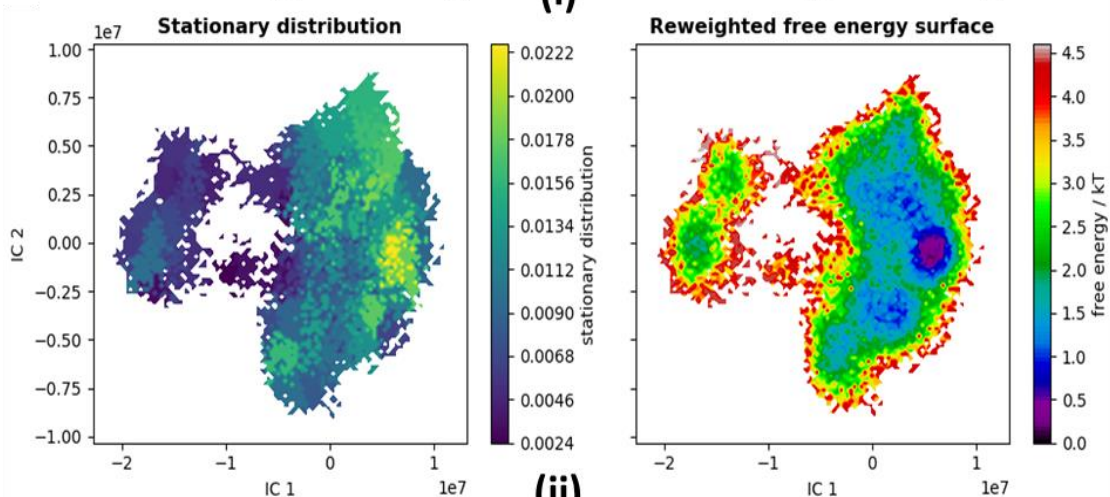
Figure 3.7B) Three kinetic equilibrium transition states have catalytic and regulatory spine mechanism while bound to cabozantinib in different states of TAM RTK kinase domain from 1 μ s MD simulations.

8A

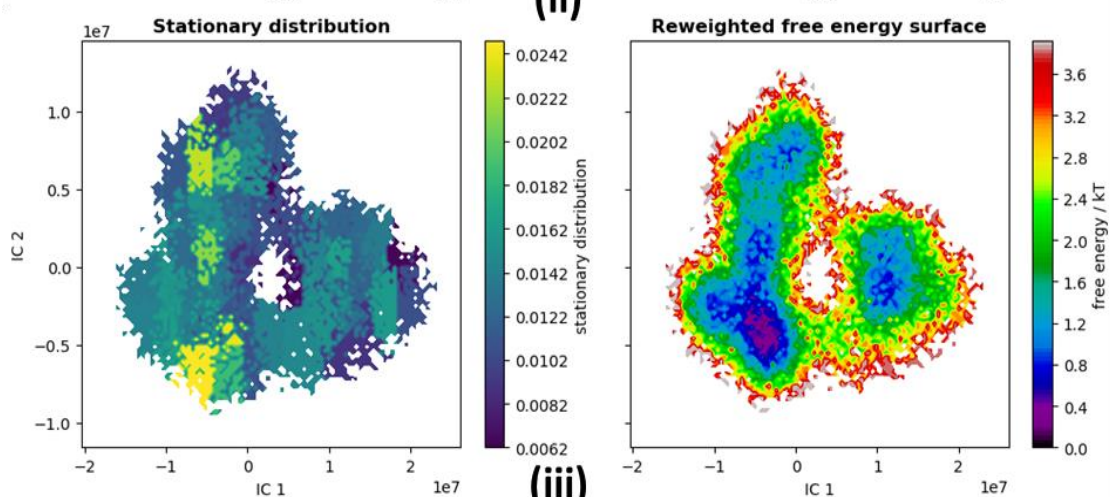
Axl-active β -sheet



(i)



(ii)



(iii)

Tyro3-active Loop

Mer-Inactive α -helix

8B

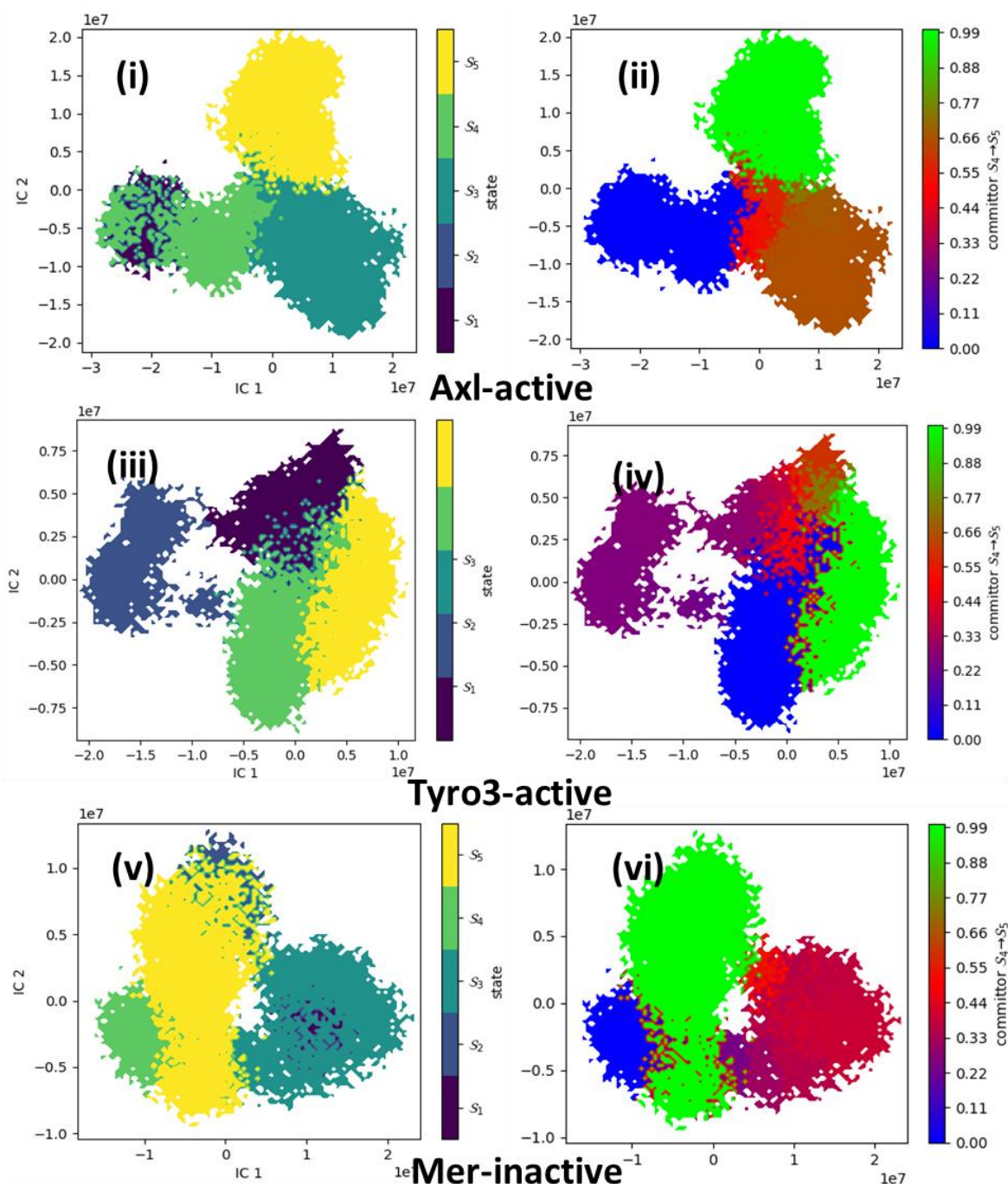


Figure 3.8) The selective kinetic equilibrium transition state models among nine hidden Markov state models of TAM RTKs.

(3.8A) Stationary distribution and free energy surface analysis of three kinetic equilibrium transition state models. (i) Axl-active; (ii) Tyro3-active; (iii) Mer-inactive.

(3.8B) Kinetic metastable five states estimation and Kinetic metastable transition state analysis three kinetic equilibrium transition state pairs. [All kinetic TS: S_4 - S_5].

From the PCA, it is revealed that all TAM kinase domains have random distribution of states. The analysis of metastable kinetic states revealed higher numbers of active state models in Axl and Mer kinase domain than the number of kinetic transition states of inactive forms, however Tyro3 kinase domain has similar numbers of coexisted metastable state transitions among the active and inactive forms. The MSMs (Husic and Pande, 2018) of the TAM kinases is expressed as different relaxation timescale intervals. The Tyro3 apo, Axl inactive and Mer active have higher relaxation timescales. The Tyro3 has equal contribution in active and inactive stationary distributions among kinetic HMM states. The five state kinetic metastable models were designed on the basis of active space distribution of HMMs of TAM RTKs. The apo Axl, inactive Mer and inactive Tyro3 have higher transition energies. These kinetic states are further validated with five MSM systems to emphasize the hidden markov dynamic state analysis. Among the nine kinetic metastable states, three HMM states are classified as “non-equilibrium kinetic transition states” (Figure 3.6) (Tyro3-apo S_3 - S_5 , Axl-inactive S_2 - S_5 and Mer-active S_2 - S_4) due to different kinetic transitions occurring among them (Table-3.1). The rest of the six HMM states undergo S_4 - S_5 kinetic transitions among five state model system mentioned as “equilibrium kinetic transition states” (Figure 3.8). The activation loop undergoes β -sheet formation in the case of active Axl and α C-helix formation in the case Mer inactive state during S_4 - S_5 kinetic transitions. In the case of Tyro3 active and inactive states, the activation loop remains in a random loop conformation. The TAM receptor tyrosine kinase bound with ATP as active state mode to facilitate phosphorylation of substrate (tyrosine amino acid). But the inhibitor (cabozantinib) bound to TAM RTKs active and inactive mode states. This result describes that the effective inhibitor bound the active receptor tyrosine kinase to arrest the substrate bound state of kinase domain to effectively block overexpressed TAM RTK's. The inactive state kinase bound inhibitor could arrest the change of protein conformations in signal transduction process to initiate the effective apoptotic signals to nullify the any malignant protein bound inhibitor state by immune cells. Therefore, the kinase bound specific states are very crucial to understand the RTKs involved in various types of cancers. In summary, salient changes in the spatial conformational states due to inhibitor binding to the active site during MD simulations in various regions of Tyro3, Axl and Mer kinases were observed. From these research findings, the kinetic active and inactive state mechanisms could explain how cabozantinib arrests the overexpressed TAM RTKs in malignant cells, a key step to inhibit the kinase signaling pathway in cellular signaling process.

3.4. Conclusions

From one μ s MD simulations each of apo, ATP, cabozantinib bound active and inactive TAM kinases, metastable active and inactive conformational states are revealed. The α C-helix region is highly distinguished and its conformational flexibility is complementary to the activation loop. The dynamical movement of the overall regulatory and catalytic spines consisting hydrophobic residues coordinated in kinase internal domain initiate cellular signal transduction process. The regulatory spine is intact and vertically aligned in ATP bound TAM kinases that continue to remain in the active conformation. However, it is broken in cabozantinib bound active and inactive TAM kinases due to the expansion of protein core arising from fluctuations in P-loop, α C-helix and activation loop. The selective TAM kinase inhibitor (cabozantinib) arrests the overexpression of kinase domains via blockage of dynamical movement of both these spines by undergoing a fragmentation of hydrophobic surface at binding site between α C-helix and activation loop (shown in Figure 3.2). The RTKs bind inhibitor in two different conformations, as active and inactive states, (K567 P-loop- α C-helix E-584 inward/outward rotations in the case of Axl). The kinase activation is in the active state mode, as the distance plots show the active states of Tyro3, Axl and Mer have large core expansion between α C-helix- activation loop in regulatory active site (Figure 3.5). The cabozantinib binding stabilised the hidden Markov state structures of active and inactive Axl, whereas the hidden Markov state conformations from the three Mer structures are closely associated with each other. From PCA, it is revealed that all TAM RTK kinase domains have random distribution of states. The analysis of metastable kinetic state forms revealed higher numbers of active state models in Axl and Mer RTKs than the number of kinetic transition states of inactive forms, however Tyro3 RTK has similar numbers of coexisted metastable state transitions among the active and inactive forms. The MSMs of the TAM kinases is expressed as different relaxation timescale intervals. Three HMM states are classified as non-equilibrium kinetic transition states (Tyro3-apo S_3 - S_5 , Axl-inactive S_2 - S_5 and Mer-active S_2 - S_4) due to different kinetic transitions among the nine kinetic metastable states (Table-3.1). In the presence of inhibitor, kinase domain proceeds as inactive state to block transduction of cellular mechanistic signal pathways in cancer therapy. The one μ s MD simulations each of apo, ATP and cabozantinib inhibitor bound active and inactive TAM kinases describes the abnormal activation and overexpression of RTKs resulting in several forms of cancers and inhibition.

Chapter - 4

In silico design, modelling and molecular mechanisms of Axl receptor tyrosine kinase inhibitors

Abstract

The RTKs are class of cell surface plasma membrane bound protein kinases. Kinase domain regulates multicellular communications among many immunological and regular cellular metabolic activities. However, some of the malignant cells have upregulated and overexpressed RTKs which are responsible for angiogenesis in many metastatic cancers. Axl RTK is one member of the RTK group and it is present in most of the eukaryotic cells. All metastatic cancer cells have overexpressed Axl tyrosine kinase to trigger uncontrolled growth and angiogenesis in the malignant cells. The upregulated kinases can be inhibited in its active and inactive states in the presence of small organic molecule inhibitors. Kinase inhibitors have been discovered to arrest the signal transduction pathways in the malignant cells as a therapy and cure for cancer. In this chapter, small molecule databases were screened using the pharmacophore features of macrocyclic inhibitor (7YS) taken as reference from the crystal structure of Axl kinase domain. Pharmacophore based virtual screening of small molecule libraries (ChEMBL32, ChemDiv, ChEMspace, Mcule, MolProt, PubChem and Zinc), followed by molecular docking, MD simulations and trajectory analysis using PCA was studied. The molecular basis for the binding of macrocyclic inhibitor and seven screened hit molecules bound at Axl kinase domain in two different modes at catalytic and regulatory sites was analysed. The specific selective inhibition of the Axl RTK by small molecules inhibitors will be the future direction study.

4.1. Introduction

TAM RTKs are activated by Gas-6, Pros1 and EGF extracellular factors. It has been proposed that TAM kinases are emerging as a class of innate immune checkpoints that participate in key steps of anti-tumoral immunity (Akalu, Rothlin, & Ghosh, 2017). TAM RTKs are also reported to play crucial roles in disease conditions such as acute myeloid leukaemia, breast, colorectal, lung, ovarian cancers and glioblastoma (Rankin, & Giaccia, 2016; Knubel et al., 2014). Since TAMs are overexpressed in many oncogenic cells, some previously reported kinase inhibitors are experimentally validated as TAM RTK inhibitors (Wu et al., 2017; Wu et al., 2018). AXL RTK is proposed as a promising drug target for anti-cancer therapy (Zhu, Wei, & Wei, 2019). Despite the huge potential of TAM kinases as drug targets, fewer attempts have been made towards the design specific inhibitors to these proteins. Inhibitors validated on other kinases have been tested for their activity on Axl kinase. For example, cabozantinib, a c-MET and vascular EGFR kinase inhibitor was also shown to inhibit Axl kinase with 7 nM inhibition (Tridente et al., 2017). Cabozantinib has been approved by FDA for differentiated thyroid cancer in September 2021. “BMS-777607”, also a MET inhibitor was shown to be more active on Axl kinase (Dai, Bae, Pampo, & Siemann., 2012). Studies on some Axl specific inhibitors such as SGI-7079, TP-0903, BGB324, DP3975 and NA80xl are in progress. From ligand-based pharmacophore screening and MD simulations studies Axl kinase domain inhibitors are reported (Nagamalla et al., 2022). Considering the importance of Axl kinase as an anti-cancer drug target, computational methods have been employed to design its inhibitors using pharmacophore-based virtual screening of large chemical libraries and molecular docking. Some of these molecules are confirmed as probable hit molecules using MD simulations and post-MD data analyses.

4.2. Materials and methods

Protein structure and binding site:

The 3D protein structure of Axl kinase domain binding with a macrocyclic inhibitor at the hinge region, present in inactive state (PDB ID: 5U6B) A chain) was considered for the computational studies. The missing residues were constructed as discussed in chapter 2.

4.2.1. Pharmacophore featured virtual screening of chemical libraries against Axl kinase domain

Based on macrocyclic inhibitor (7YS) binding position to the Axl kinase domain, “Pharmit” (<http://pharmit.csb.pitt.edu>) (Sunseri et al., 2016), a pharmacophore generating online webserver was used to generate the best pharmacophore for the design of hit molecules binding to Axl kinase domain. The possible features available in the Pharmit server are hydrogen bond donor, hydrogen bond acceptor, hydrophobic and ring aromatic. The selection of pharmacophore features was based on the complementarity of Axl kinase – 7YS complex. The databases available in the Pharmit server were used for pharmacophore based virtual screening. It screened libraries of millions of compounds based upon

best pharmacophore and molecular shape modalities with advanced pre-built structural screening algorithms through the online webserver. Finally, it provides comprehensive query optimized hit molecules with online platform for structure based virtual screening method. The molecules obtained based on lower RMSD and fewer rotatable bonds were selected. An in-house database of the downloaded molecules was prepared after adding hydrogens and their structure refinement using Discovery Studio 3.5. The identified molecules were passed through the virtual screening protocol using PyRx (Dallakyan et al., 2015) by docking the molecules into the binding site of 7YS in 5U6B A chain. The screened-in molecules above a certain binding energy were selected for molecular docking studies.

4.2.2. Molecular docking of macromolecule (7YS) and screened-in molecules

The screened-in hit molecules were docked into the macrocyclic inhibitor and ATP binding pocket of the inactive state of Axl RTK using AutoDock tools (Morris et al., 2009). The macrocyclic inhibitor was initially docked to validate the docking methodology. A grid box was set up within 5 Å space around inhibitor binding that encompasses active site residues of Axl kinase and 50 docking poses were selected for each docked molecule. A grid box with the dimensions of X: 50.960, Y: 55.553, Z: 45.935, with a grid spacing of 0.508 Å was used. The docked pose of molecules with best fitting and better binding energy were selected for 250 ns MD simulations to confirm the binding pose in the kinase active site and to decipher the molecular basis for interactions with Axl kinase domain.

4.2.3. ADME data analysis of seven hit molecules

The pharmacokinetic properties like, absorption, distribution, metabolism and excretion (ADME) of the hit molecules were assessed with SwissADME (Daina, Michielin, & Zoete., 2017) web-based server tool (<http://www.swissadme.ch>). The drug-likeness properties are represented in the form of TPSA, consensus Log $P_{o/w}$, LogS (ESOL), LogKp (skin permeation) (cm/s), synthetic accessibility were also calculated (Daina, Michielin, & Zoete., 2014, Ertl, Rohde, & Selzer., 2000, Potts, & Guy., 1992).

4.2.4. Molecular dynamics simulations

All MD simulations were achieved using AMBER (Gotz, et al., 2012) version 18.14 for the seven screened hit molecules and macrocyclic inhibitor bound to Axl kinase domain. The best docking pose of each complex was utilized as input for MD simulations. The force fields for the entire systems were generated with Antechamber using am1bcc method (Wang et al., 2006, Colovos and Yeates, 1993). The input parameter files for MD simulations were generated after adding hydrogen atoms in tLEaP module in AMBER tools (Anandakrishnan, Aguilar and Onufriev., 2012, Lindorff-Larsen et al., 2010). Sodium and chloride ions were added to the systems to neutralize the charge, each molecular system was solvated within an 8 Å size cubic box. The final ionic concentration for the systems was set to 100 mM. The Amberff99sb-ILDN force field was used for entire model system with TIP3P water model for AMBER molecular parameters (Meagher et al., 2003, Mark and Nilsson, 2001), temperature was set to 300 K and 1 atm pressure with Monte Carlo barostat (Salomon-Ferrer et al., 2013). Energy minimization was carried

out by using steepest descent method for 40,000 cycles to overcome short range null contacts among the molecular system in solvent (Darden et al., 1993). Long range electrostatic interactions were considered with Particle Mesh Ewald algorithm (Jorgensen et al., 1983) with cut-off range 9 Å and order 4. All model systems were equilibrated for 7 ns before the production run, and the coordinates in the production run were saved after every 10 ps (Salomon-Ferrer et al., 2013, McGibbon et al., 2015). The MD simulations of each molecular system was carried out for 250 ns. Axl kinase in complex with ATP was also studied for comparison.

4.2.5. MD simulations data analysis

The AMBER MD trajectory data analysis was carried out using cpptraj with Amber tools 18 (Hornak et al., 2006). The average structures after MD simulations, RMSD, RMSF and specific regions were analysed with parmed and cpptraj module in AMBER18 tools. MM-GBSA and MM-PBSA calculations were carried out on 2.5k conformers from 25k frames of 250 ns MD simulations data in AMBER MMPBSA.py module. It is an efficient and user accessible end-state free energy calculating algorithm which are inbuilt in AMBER18 tools. The end-state implicit solvent model free energy calculations are divided to include Generalized-Born (GB) and Poisson-Boltzmann (PB). The free energy calculations were done with single trajectory protocol by considering 2.5k frames from MD simulations of the bound and unbound states between receptor and ligand complexes (Miller et al, 2012 and Wang et al, 2017).

4.2.6. Principal component analysis

PCA is an efficient multivariate statistical technique to extract the conformational changes in a protein with respect to time. This method reduces the number of dimensions in a systematic manner therefore, motions from largest to smallest spatial scales can be observed. The large scale motions mask the small scale motions due to their greater atomic displacements. However, the structural dynamics of biological relevance are calculated as PCA modes for a small number of motions with large amplitudes. A covariance matrix is constructed from the Cartesian coordinates of the trajectory conformations where the eigenvalue decomposition of the covariance matrix leads to a complete set of orthogonal collective modes (eigenvectors). Therefore, each eigenvalue contributes to a small portion of the motion. These eigenvalues were obtained by the diagonalization of the covariance matrix to the C α atomic fluctuations and it follows the decreasing order of the corresponding eigenvector indices. The square matrices were diagonalized into secular equations to generate eigenvalues and corresponding eigenvectors. These eigenvectors are the best principal components associated with the respective eigenvalues (Yang, Eyal, Bahar, & Kitao, 2009).

4.3. Results and discussions

In the crystal structure of Axl kinase (5U6B), Asp585 side chain is pointing inwards and makes ionic interaction with Lys567 in the B chain, that is indicative of a kinase in an active state. Whereas, the

Asp585 side chain is pointing outwards and therefore does not make ionic interactions with Lys567 in the A chain that is indicative of an inactive state. In both the chains, Phe691 aromatic side chain is facing inwards into the hydrophobic cavity formed by the side chains of amino acids Met589, Phe592, Met598, Phe668, Leu663, Val597 and Val688. The crystal structure of Axl kinase in the inactive state (5U6B, A chain) was selected for computational studies. The macrocyclic inhibitor, 7YS makes hydrogen bonds with the protein. The main chain NH of Met623 forms hydrogen bond with the pyrazine nitrogen, the Asp627 NH and the side chain carboxylate group form hydrogen bond with the terminal OH group of 7YS. Pro621 acts as a hydrogen bond acceptor with primary amine on the pyrazine. The regulatory spine consisting of four non-consecutive hydrophobic amino acid residues; Leu600 (β_4 -strand); Met589 (α -helix); Phe691 (DFG motif); His670 (catalytic loop) and an additional residue Asp731 from the C-terminal lobe align vertically from N - terminus to the C -terminus.

4.3.1. Pharmacophore model generation for virtual screening

Based on the binding of macrocyclic inhibitor- 7YS, the best pharmacophore was generated. Hydrogen bond donor, hydrogen bond acceptor and aromatic features were selected as pharmacophore features. The four hydrogen bonds and three aromatic rings on pyrazine, phenyl and pyrazole were considered as pharmacophore features on 7YS. These features truly represent the biological features required for the complementarity between the Axl kinase and potential inhibitors. Among the molecules obtained from pharmacophore-based screening of databases, the molecules that display lower than 0.7 Å RMSD and fewer than 6 rotatable bonds were selected. The virtual screening of ChEMBL32 small molecule database identified (2,186,411 molecules having 28,970,382 conformations identified hits 75 hits), ChemDiv (1,456,120; 21,462,597 and 9 hits), ChEMBL (50,181,678; 250, 205, 463 and 3 hits), Molecule (45, 257, 086; 223,460,579 and 15 hits), MolPort (4,843,718; 67,033, 884 and 27 hits), PubChem (103,302,052; 499,442,812 and 1,199 hits) and Zinc (13,127,550; 122,276,899 and 65 hits). The selected molecules were proceeded for virtual screening using PyRx, that identified 30 molecules with a score ≤ -7.5 kcal/mol to bind the active site of Axl kinase.

4.3.2. Molecular docking of screened molecules

The 30 hit molecules selected from virtual screening were docked into the macrocyclic inhibitor binding site using AutoDock tools. The molecules having binding energy with ≤ -6.5 kcal/mol and that make hydrogen bonds with active site of the receptor were selected for further MD simulations. The AutoDock grid parameters, docking binding energies and molecular weights of hit molecules along with 7YS are provided in Table 4.1. Based on molecular docking, it is observed that the hit molecules bound in a related manner to 7YS, so that Axl RTK can be inhibited with these hit molecules that possess similar pharmacophore features from chemical and steric insights. From the overall docking studies, the pose with highest docking scores of seven hit molecules in complex with Axl kinase domain were selected.

	Axl kinase-ligand complex	Docking Axl-inhibitor mean binding energy in kcal/mol	Docking Grid parameters (x, y, z)	Molecular weight (g/mole)
01	Hit-1	-8.01	(54.876, 57.627, 56.255)	415.53
02	Hit-2	-7.65	(54.876, 65.126, 54.550)	438.33
03	Hit-3	-6.64	(54.876, 61.837, 52.935)	440.44
04	Hit-6	-7.40	(54.910, 61.830, 57.477)	411.84
05	Hit-4	-6.80	(52.960, 60.553, 52.935)	369.48
06	Hit-5	-8.29	(54.883, 60.915, 55.551)	412.51
07	Hit-7	-6.96	(54.883, 61.514, 56.388)	396.44
08	7YS active /inactive state	-7.15	(50.960, 55.553, 45.935)	460.89
09	7YS inactive active state	-6.75	(50.960, 55.553, 45.935)	460.89

Table 4.1) Docking parameters of seven screened hit molecules with reference 7YS macrocyclic inhibitor.

4.3.3. ADME data analysis and docking data correlations

The synthetic accessibility of all hit molecules range from 3.3 to 5.8, this is indicative of the ease of their synthesis with current synthetic methods. The TPSA represents the ability of the molecules to cross the biological barrier for tissue absorption and their brain access in our body and most hit molecules are within the reasonable range. Lipophilicity is the parameter to assess the partition of the molecules between n-octanol and water [$P_{o/w}$]. This value ranges between 0.5 - 1.9 indicative of their effective partitioning. LogS indicates the water solubility of probable ligands, the seven hit molecules are comparable with reference molecule (7YS) (-3.62). It indicates the ease of handling drug formulation in pharmacodynamics of drugs in clinical trials. The Log K_p value indicates the linear regression of molecular size and lipophilicity. The lower skin permeant of molecule has more of negative value of Log K_p . The Log K_p values above -7.0 indicate that the molecules are less permeable through the skin. Based upon ADME data, the drug likeness for screened hit molecules can be considered. These are the key parameters to be studied for development for selective hit molecules for their physiochemical properties to be approved as oral drug candidates. These parameters are provided in Table 4.2.

S. No	Axl kinase-ligand complex	TPSA (Å ²)	Consensus Log P _{o/w}	LogS (ESOL)	Log Kp (skin permeation) (cm/s)	Synthetic accessibility
01	7YS	119.39	1.92	-3.62	-8.37	5.05
02	Hit-1	83.66	1.62	-3.21	-7.59	5.77
03	Hit-2	89.96	1.45	-4.10	-7.06	3.48
04	Hit-3	84.86	1.80	-3.53	-7.56	4.02
05	Hit-6	144.03	1.27	-2.67	-8.72	3.32
06	Hit-4	95.92	0.57	-2.01	-8.33	3.94
07	Hit-5	122.77	1.59	-3.47	-7.57	4.09
08	Hit-7	99.69	0.96	-2.70	-8.00	4.52



Table 4.2) The ADME properties of the hit and reference (7YS) molecules

Standard values for reference TPSA= 0 to 140 Å², Log Po/w= −4.0 to 5.6, Log Kp= −6.1 to −0.19 cm/s, Synthetic accessibility scale= 1 to 10. LogS, solubility of molecules in aqueous medium; 0 to −2 are soluble, those in the range of −2 to −4 are slightly soluble and insoluble if less than −4.

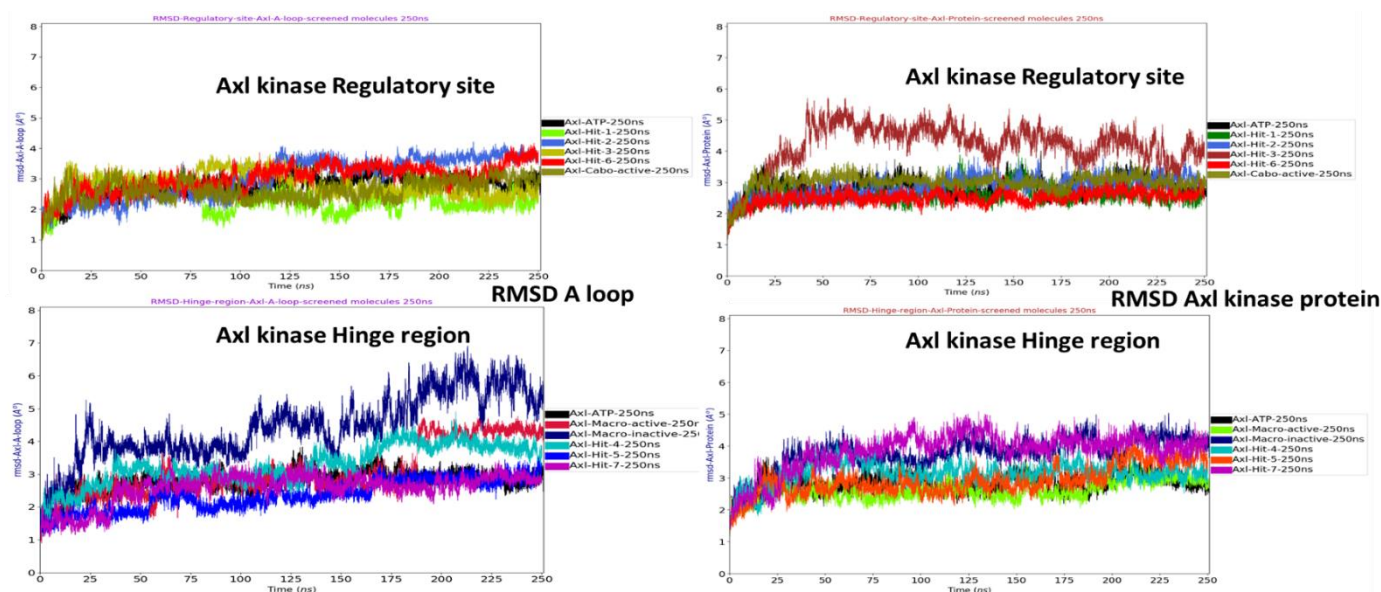


Figure 4.1A) RMSD plots of Axl RTKs (protein and activation loop) complexed with seven hit molecules in active and inactive states from 250 ns MD simulations.

4.3.4. Molecular dynamics simulations

The 250 ns MD simulations were carried out using AMBER on inactive Axl kinase complex with macrocyclic inhibitor, 7YS and the screened-in hit molecules. This would reveal how the macrocyclic inhibitor and hit molecules differentiate between the active and inactive states and the various conformational states in between, in order to influence the structural features of various regions in its 3D structure. This can be revealed by RMSD, RMSF and regulatory spine analysis. The hit molecules bind the Axl kinase in a stable manner throughout the simulations. The intermolecular hydrogen bonding indicate the stable complex formation. An inhibitor can bind the kinase domain at the catalytic site, where the ATP molecule also binds involving the hinge region and is called as “catalytic site of substrate”. The regular phosphorylation takes place on hydroxyl side chain of tyrosine present in the target protein by RTK, and referred to as “regulatory site for target substrate”. Based upon this, the inhibitors bound to kinase domain are site specific while inhibiting the kinase structural conformational changes occurring during effective cancer therapy. Therefore, the screened hit molecules induced structural deformations in kinase domain based upon their site selective inhibition in Axl kinase domain. Based on the MD simulations trajectories it is observed that the seven hit molecules can be classified into two binding modes; four screened molecules (hit 1, 2, 3 and 6) bind at regulatory site region near the junction of P-loop, α C-helix and activation loop; and three hit molecules bind the hinge region (hit 4, 5 and 7) of Axl kinase domain. The complexes with hit 1, 2, 3 and 6 have lower RMSD when the initial and average structures are superposed as shown in Figure 4.1. In the complexes, 4, 5 and 7 the RMSD is higher and the α C-helix has moved considerably compared to the initial structures and largely it has been observed that the N-terminal β -sheet domain has deviated. The macrocyclic inhibitor bound Axl kinase is hinge region bound, but has a low α C-helix deviation compared to the initial structure. It can be seen from the RMSD plots that the C α atom fluctuations are relatively lower in the regulatory region bound hit molecules rather than hinge region bound hit molecules with the exception of the hit 3 (4.5 Å). The higher RMSD in the Axl - hit 3 complex is due to the kinase domain that entirely exists in the inactive state. In the 250 ns MD simulations trajectory showing 12.3 Å distance between K567-E585 is shown in the Figure 4.2A. The Axl bound hit 5 continues to exist as inactive conformer during 250 ns MD simulations, however the hit 4 and hit 7 complexes are stabilized in the active state as can be seen from the MD simulations. The salt bridge (K567-E585) distance between P-loop and α C-helix regions are below 3.0 - 4.5 Å for active states (hit 4- active, hit 7- active), and hit 5-inactive salt bridge distance varies between 3.0 -17.5 Å (Figure 4.2B). The hit 4 complexed Axl kinase has higher RMSD in the activation loop compared to the other hit molecules. The Axl kinase bound to 7YS has a larger activation loop dynamics (Figure 4.1). On the whole, the Axl- bound hit-3 (inactive state, regulatory site) and hit-7 (active state, hinge region) display higher RMSD values among all the hit molecules. Surprisingly, the RMSF plots of

all seven hit molecules superpose well due to similar influence of all side chains in Axl kinase domains irrespective of the site selective inhibitions as shown in Figure 4.3. The regulatory site binding hit molecules are located near to the DFG motif and activation loop region thus disrupting the regulatory spine between Met589 (α C-helix) - Phe691 (DFG) motif activation loop region, leading to changes in the overall kinase domain structure. The broken regulatory spine can be seen from the Figure 4.4A.

4.3.4.1. Axl bound ATP and the inactive/active Axl kinases complexed with 7YS

The ATP cofactor binds to Axl kinase domain catalytic region that connects between kinase N-terminal and C-terminal domain with hinge loop region. ATP bound at Axl catalytic domain only influences the hinge region residues and there is no involvement of regulatory spine containing residues from P-loop, α C-helix, activation loop (DFG motif) and catalytic loop (C-terminal domain). Therefore, the regulatory spine is retained in ATP bound Axl kinase domain. However, 7YS (macrocyclic inhibitor) bound active and inactive states of Axl kinase domain also influence the hinge region. This has also resulted in retained regulatory spine in 7YS bound to both states of Axl kinase. The stereo spatial orientation of amino acids that form these regulatory spines are altered due to the conformational changes in the Axl kinase when complexed with the four screened small molecules (hit 1, hit 2, hit 3, hit 6). The regulatory spine is retained for the Axl bound to hit 4, hit 5 and hit 7. The key components leading to the regulatory spine breakage where the inhibitors are specified to bind at regulatory site region were examined. The regulatory spine breakage mechanism is influenced by the specific homo and heterocyclic functional groups binding at the junction of P-loop, α C-helix and activation loop. The regulatory site binding inhibitors influence the Axl kinase domain regions at regulatory spine and P-loop, α C-helix and not the activation loop. This could reveal that the inhibitor binding site at specific spatial orientation of kinase plays a key role in signal transduction during kinase inhibition. The regulatory spine analysis measures the protein – ligand binding mechanisms. The binding free energies ascribe the efficacy of kinase-inhibitor integration to classify ligands as the kinase binding inhibitors. Specified cyclic groups from various screened hits small molecules are major components of regulatory spine dissociation in kinase domains. Hit 1 (adamantanyl), hit 2 (p-chlorophenyl), hit 3 (1,4 dioxane), hit 6 (5-chloro indole) are key functional groups attached to screened-in hits bound at the narrow end, away from hinge region in kinase domains. The hinge region kinase blockers are caused mainly by small molecular size to effectively fit into hinge region of the catalytic site. Five membered heterocyclic imidazole (hit 4) and saturated bicyclic pyrrole groups (hit 5, 7) mostly bound at hinge region site of Axl kinase domain (Figure 4.5)

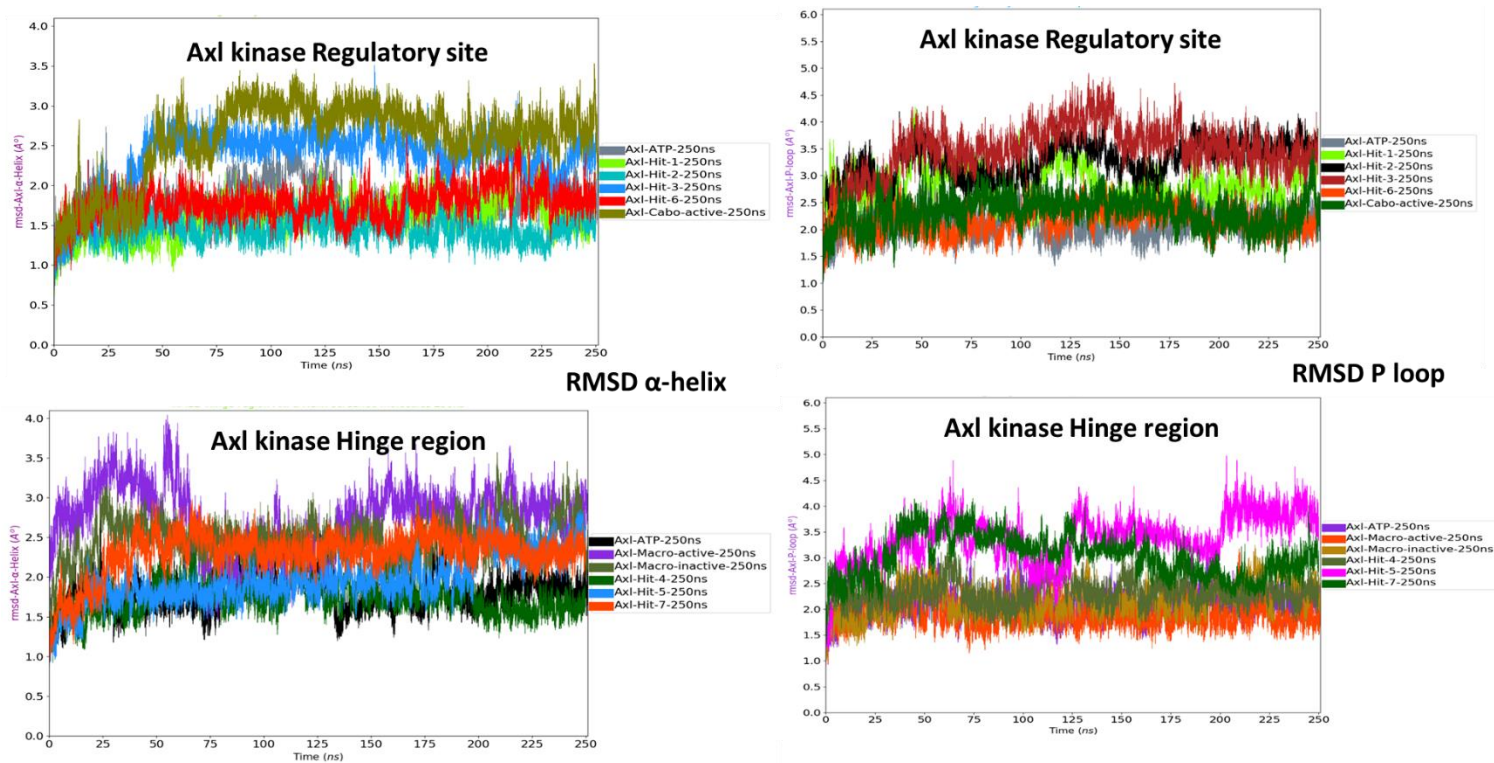


Figure 4.1B) RMSD plots of Axl RTKs (α -helix and P loop) complexed with seven molecules in active and inactive states from 250 ns MD simulations.

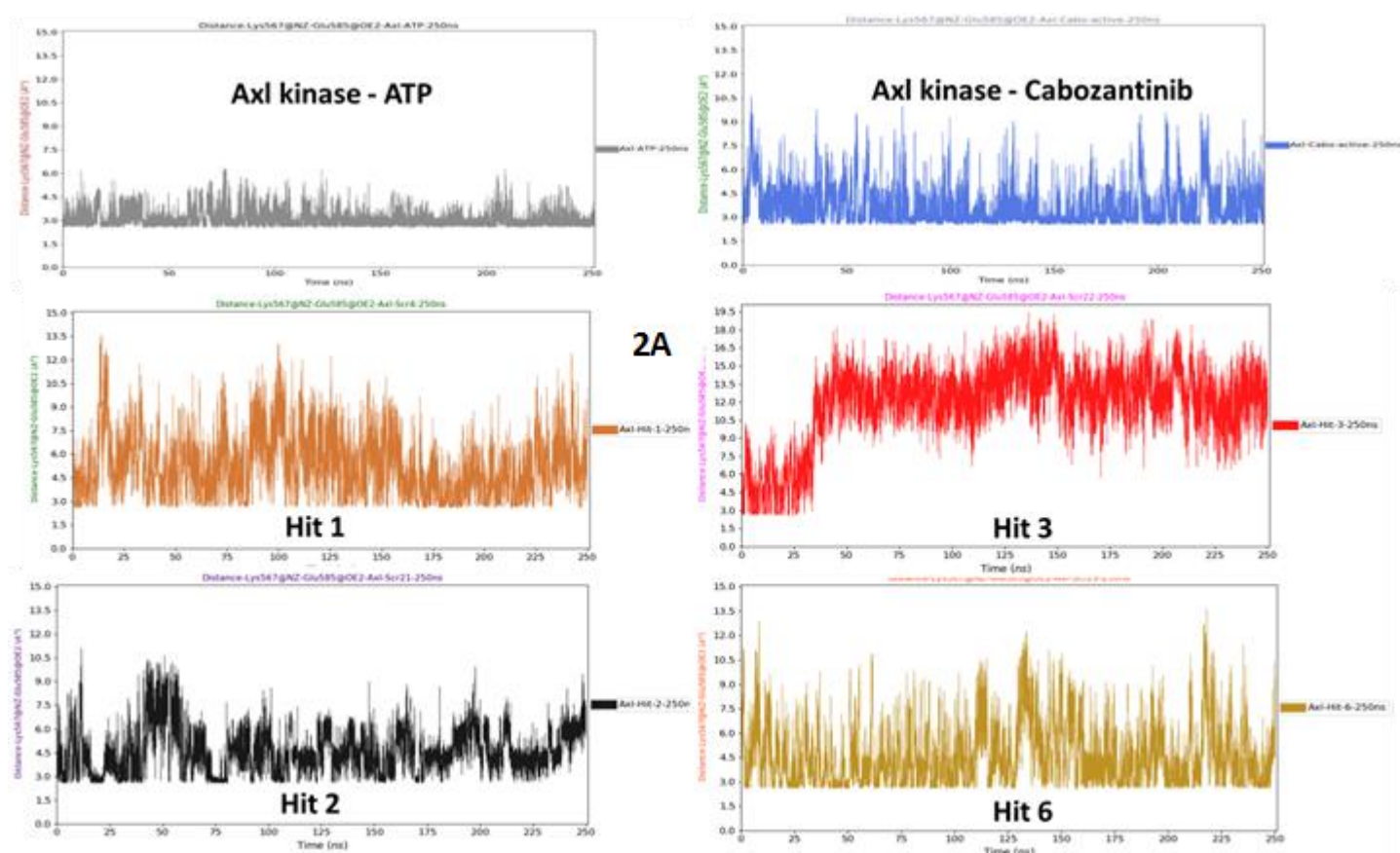


Figure 4.2A) (P-loop) Lys567 – Glu585 (α C-helix) salt bridge distance Axl kinase regulatory site bound hit 1, 2, 6 molecules bound active and hit 3 inactive states from 250 ns MD simulations.

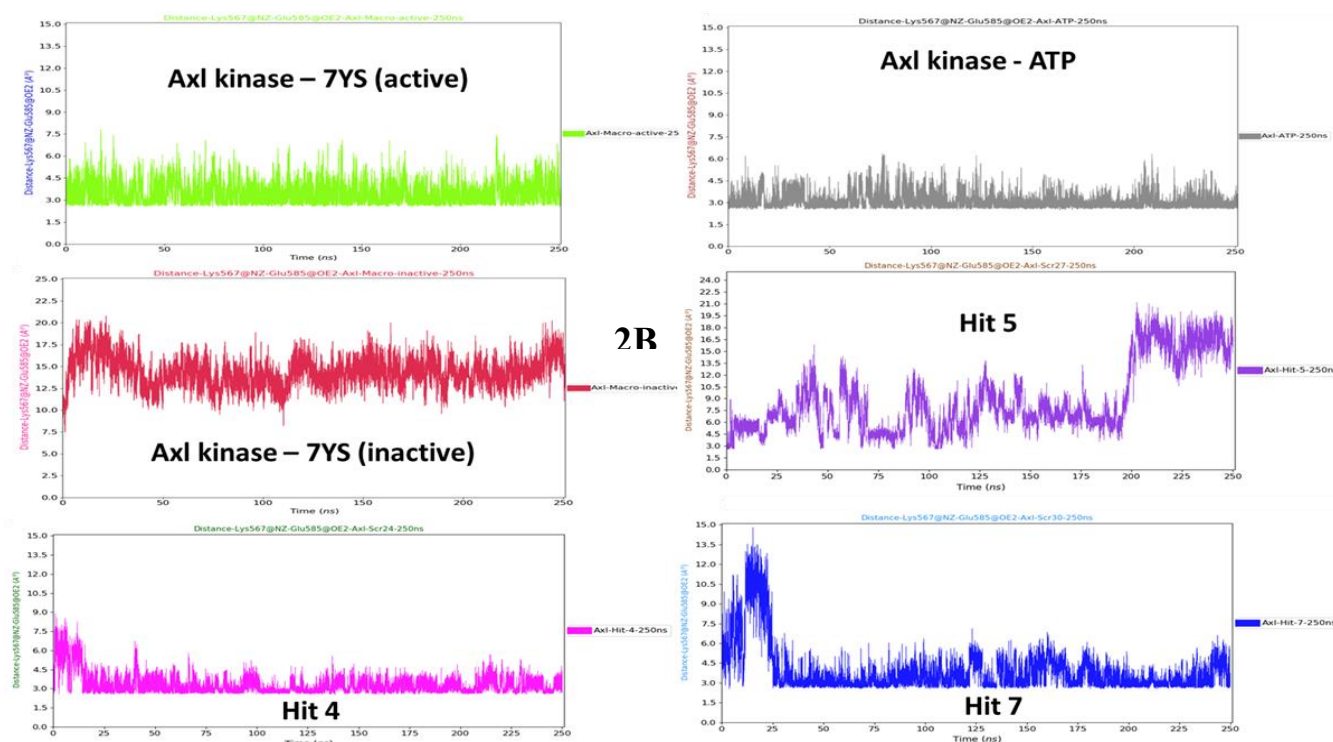


Figure 4.2B) (P-loop) Lys567 – Glu585 (α C-helix) salt bridge distance Axl kinase hinge region site bound ATP, 7YS, hit 4, 7 molecules bound active and 7YS, hit 5 inactive states

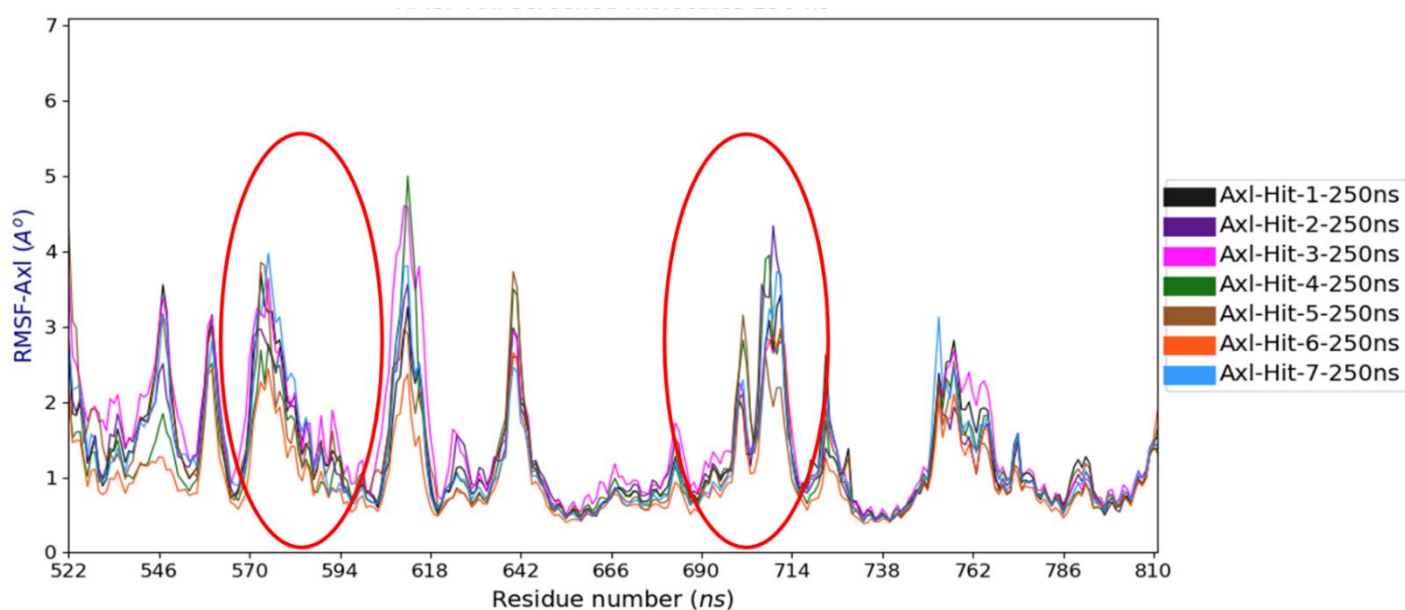
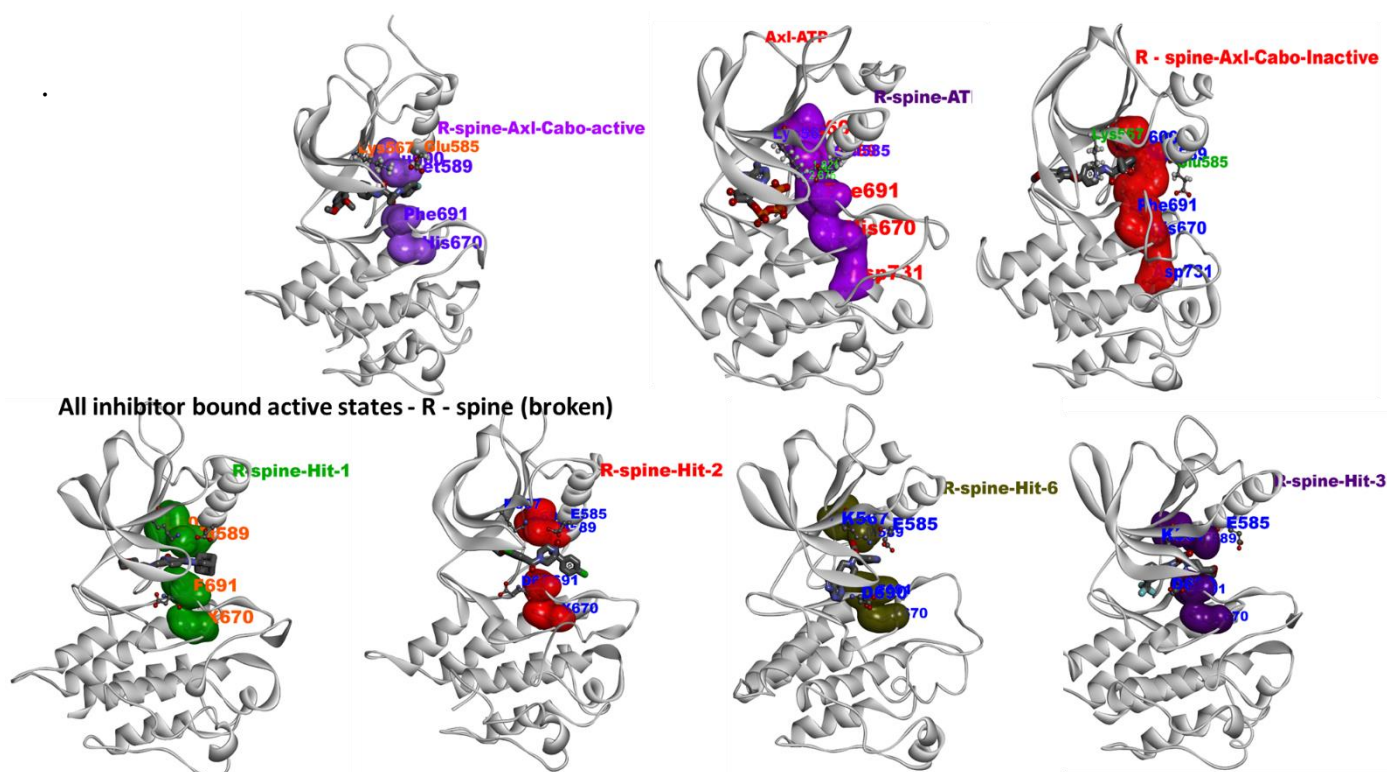
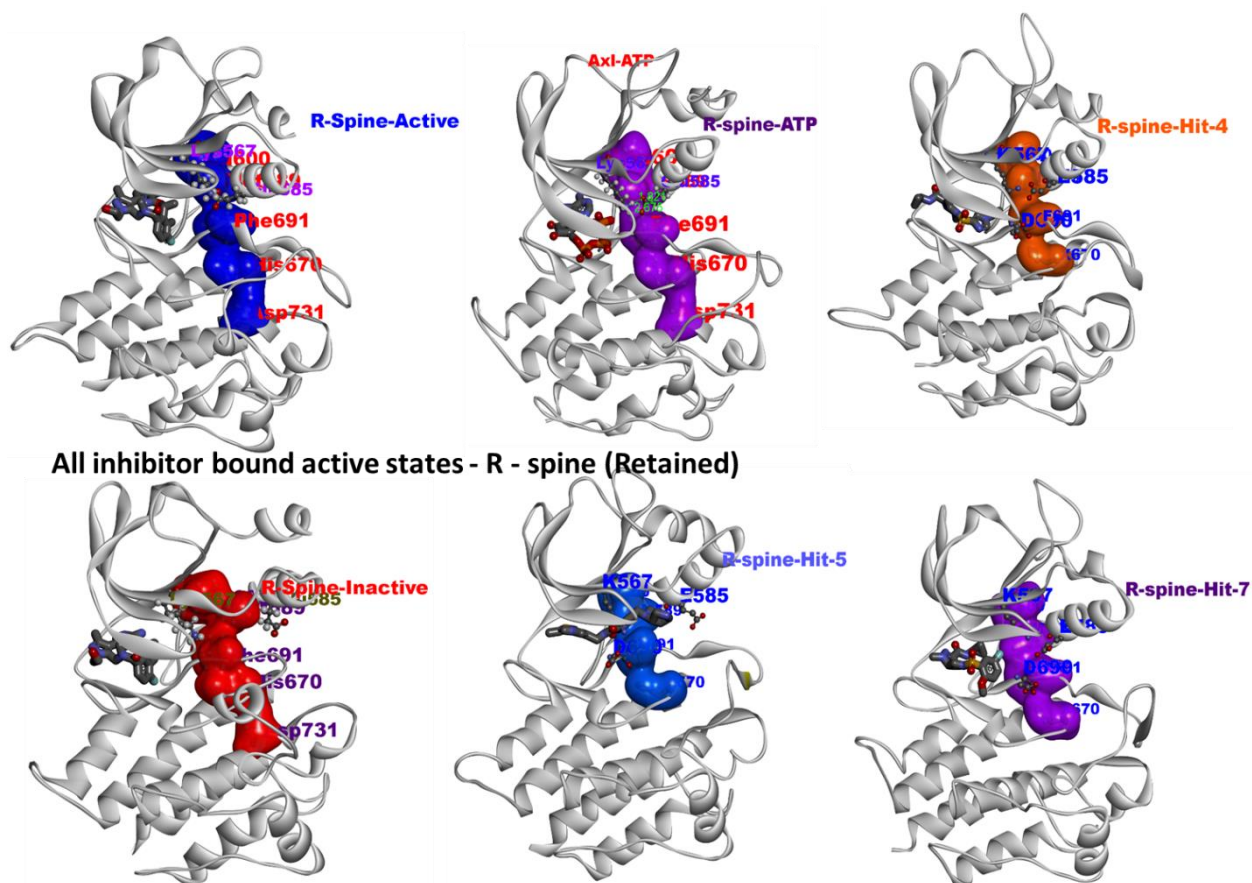


Figure 4.3) RMSF plots of Axl RTKs complexed with seven hit molecules bound active and inactive states from 250 ns MD simulations. For the sake of convenience all Axl kinases are numbered from 521 onwards. 533-565 (β turn in the N-Terminal domain); 568-598 (first α -Helix in the C-Terminal domain); 690-724 (activation loop);



4A) Axl kinase regulatory site region inhibitors



4B) Axl kinase hinge region binding site inhibitors

Figure 4.4A) Regulatory spine breakage takes place in regulatory site bound hit 1,2,3,6 molecules. **Figure 4.4B)** Regulatory spine is retained in the hinge region bound site ATP, 7YS, hit 4, 5, 7 molecules from 250 ns MD simulations.

4.3.5. Binding free energy calculations

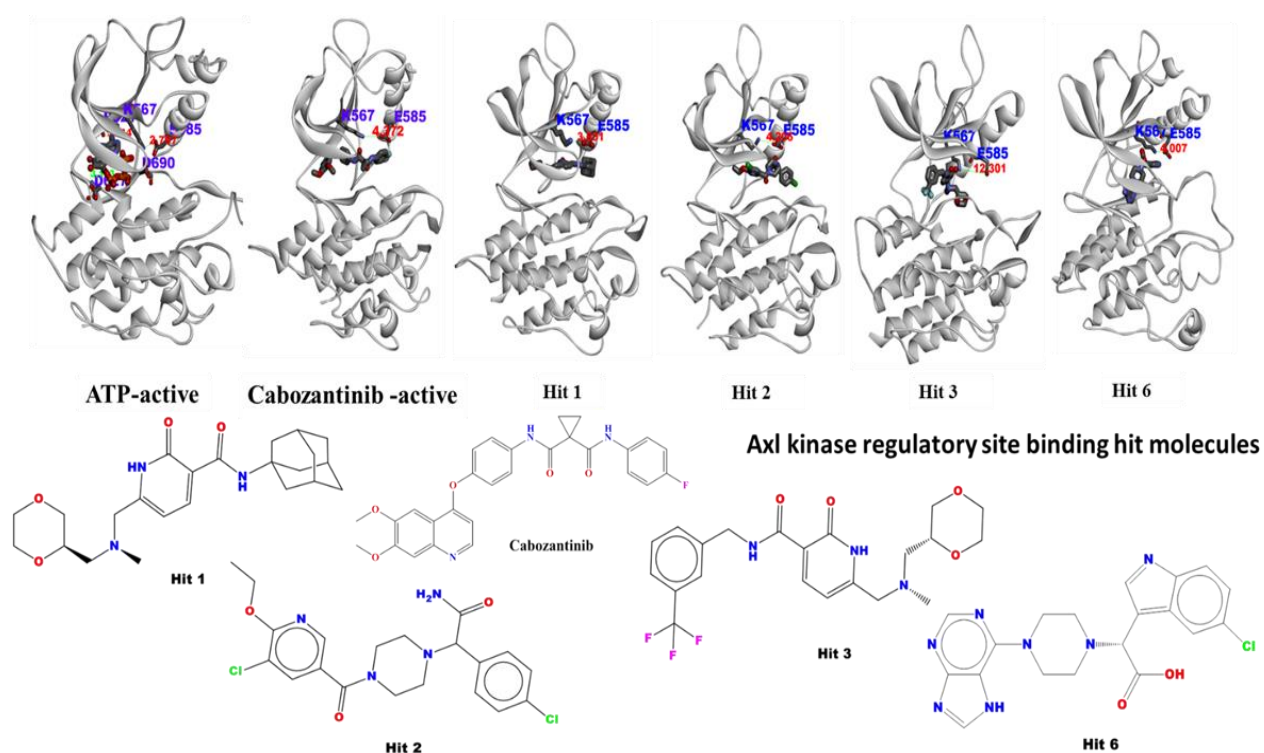
4.3.5.1. MM-GBSA (Binding affinity energy)

The data points derived from 2,500 conformations from 25k frames of trajectories with interval of 100 ps from 250 ns MD simulations data were used to calculate the binding free energies between the receptor (kinase domain) and ligand (hit molecules and 7YS) in the presence of solvent, that are shown in Table 4.3. The regulatory site kinase inhibitors binding affinity can be explained based upon binding energies in entire the MD simulations trajectories. The ATP bound kinase active state has higher binding free energy (-51.02 kcal/mol) due to its catalytic site binding and involved in the transfer of γ – phosphate group to substrate molecules in kinases enzymatic action.

Free Energy Calculation of Differences (Complex – {Receptor – Ligand})						
S.No	Protein-Inhibitor complex	GB Energy (kcal/mole)	Std. Err. of Mean	PB Energy (kcal/mole)	Std. Err. of Mean	R – spine status
	R-site region bound	Gibbs free E		Binding E		
03	Axl -active Hit-1	-42.9137	1.0784	-8.8192	0.4919	Broken
04	Axl -active Hit-2	-29.8852	1.9735	-3.9289	0.4559	Broken
05	Axl-Inactive Hit-3	-42.2415	2.4616	-14.5832	0.9805	Broken
06	Axl -active Hit-6	-35.8968	0.3405	-11.7482	1.6533	Broken
	Hinge region bound					
07	Axl-active Hit-4	-38.5077	0.1154	-10.2352	1.6673	Retained
08	Axl-Inactive Hit-5	-46.8368	0.1598	-10.7742	1.6998	Retained
09	Axl-active Hit-7	-32.5000	0.2113	-6.4635	1.6563	Retained
10	Axl - active -ATP	-51.0202	0.1737	-14.3864	1.6290	Retained
11	Axl –Macro –inactive	-37.0225	0.2116	-9.5484	1.6295	Retained
12	Axl – Macro –active	-38.6647	0.1697	-9.4953	1.6191	Retained

Table 4.3) MMPB/GBSA free energy data analysis for seven screened hit molecules with references (7YS, ATP)

5A) Axl kinase regulatory site region bind inhibitors



5B) Axl kinase hinge region binding site

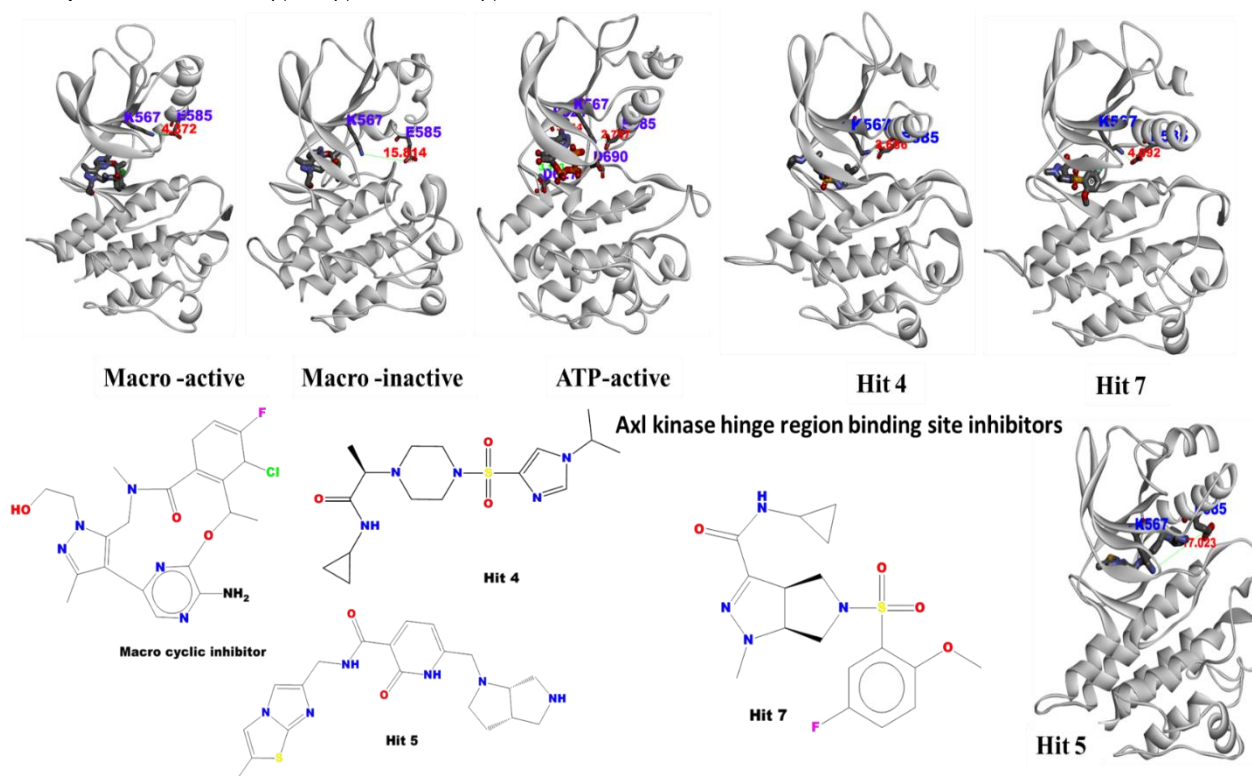


Figure 4.5A) Regulatory site bound hit 1, 2, 6 (active), hit 3 (inactive) molecules with reference of ATP. **4.5B)** Hinge region bound site ATP, 7YS, hit 4, 7 (active), hit 5 (inactive) molecules from 250 ns MD simulations.

The hinge region Axl kinase hit molecules bound active state conformer with lower binding free energies; -38.5 kcal/mole (hit 4) and -32.5 kcal/mole (hit 7) and a higher binding free energy for inactive state -46.83 kcal/mole (hit 5). The regulatory site Axl kinase hit molecules bound inactive state conformer has lower binding free energies; -42.24 kcal/mole (hit 3) and a nearly equal binding free energy for active state -42.91 kcal/mole (hit 1), -29.88 kcal/mole (hit 2) and -35.89 kcal/mol (hit 6). The binding free energies for the active Axl kinase (-38.66 kcal/mol) and inactive Axl kinase (-37.02 kcal/mol) when complexed with macrocyclic inhibitor (7YS). The kinase inhibitor binding affinity can be explained with energy values due to the overall binding energies influenced by energy terms calculated for protein-ligand in the presence of solvent waters. The ATP cofactor binding to kinase domain has highest binding energy among all kinase inhibitors and hinge region bound kinase domain inhibitors have closest binding energy correlated to ATP cofactor binding analogy. This can be further discussed with overall Gibbs free energies among kinase screened-in hit molecules.

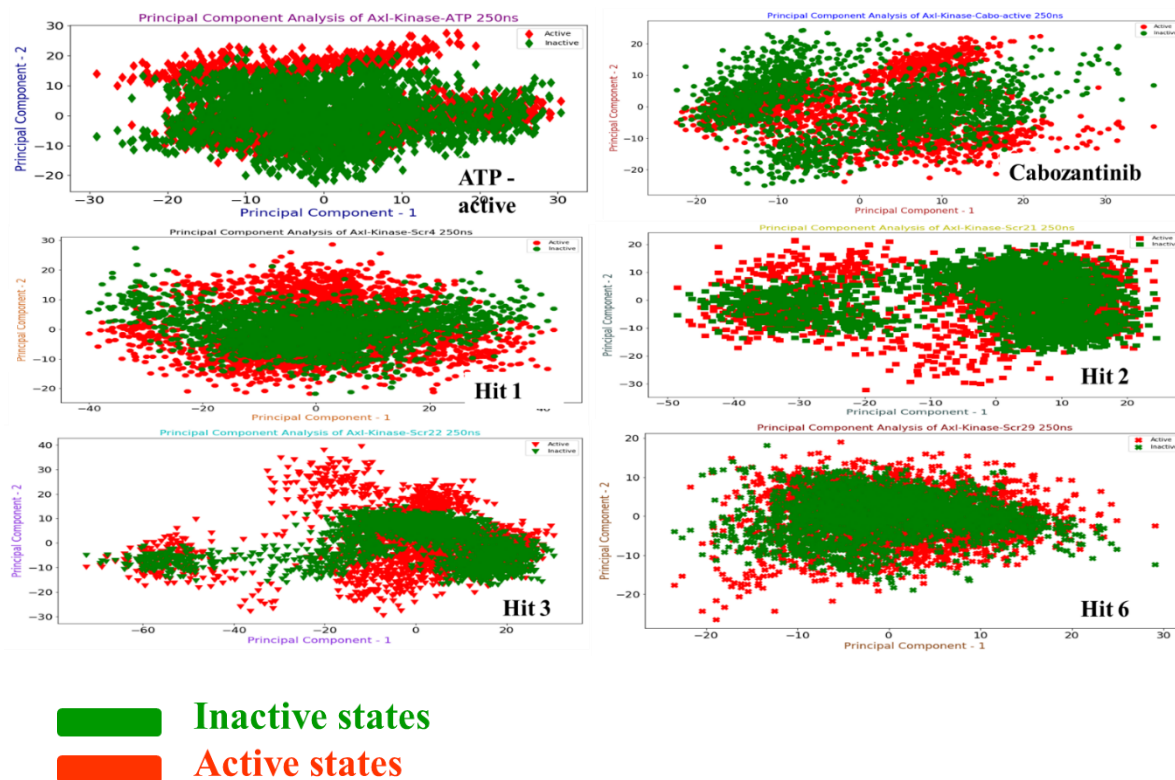
4.3.5.2. MM-PBSA (Binding Gibbs free energy)

The overall Gibbs free energies differentiate between the selective site inhibitors of kinase domain states. MM-PBSA is quite different from MM-GBSA due to the entropy energy terms calculated including the NMA of overall conformers from MD trajectories in the absence of solvent. Therefore, these energy values are lower than MM-GBSA. The regulatory site bound hit molecules have energy values varying from -8.8 kcal/mole (hit 1), -3.92 kcal/mol, (hit 2), -11.74 kcal/mole (hit 6), kcal/mole in the active state and in the case of hit 3 (-14.58 kcal/mol) that existed as inactive state and has the highest overall Gibbs free energies. This value is comparable with the ATP bound active state mode (-14.38 kcal/mol). The active and inactive Axl kinase domain bound to 7YS with similar binding affinities; -9.5 kcal/mol. The Gibbs free energies are in correspondence with the molecular docking parameters given in Table 4.3. In a macroscopic pictorial view, the overall Gibbs free energy changes are mostly accumulated into inactive states of kinases bound ligands as represented in Table 4.3. The Axl-hit-3 (-14.58 kcal/mole), Hit-5 (-10.77 kcal/mole), Axl- macrocyclic inhibitor (7YS) (-9.5 kcal/mole). The hit molecules bound to kinase inactive state are comparable with Axl-ATP. It is clearly evident that where kinase bound ATP complex is retained in active state mode of Axl kinase conformation with high affinity and binding free energy is similar to inactive states of inhibitors bound Axl kinase domain. This could be revealed as Gibbs free energetics that showed a correlation between cofactor and hit molecules bound Axl kinase states. These hidden states can be explained with PCA analyses of all the complexes in detail.

3.4.6. Principal component analysis

The MD simulations trajectories of the Axl kinase – hit molecule and 7YS complexes were analysed with PCA analysis. The regulatory site kinase bound hit molecules have large clusters of kinase inactive (green) and active (red) conformers represented in the scatter plots (Figure 4.6) and it overlapped states on each other. But the hinge region bound hit molecules are divided into two different groups of clusters represented as conformational movement towards hinge region site of Axl kinase (Axl- 7YS complex). The clustering of kinase states overlap with each other in regulatory bound inhibitors from their initial states. But the hit 4, hit 5 and hit 7 molecules are more prone to dissipate as clusters through the PCA space and are represented as mixed kinase state system in given unsupervised 2.5k conformers derived from 250 ns MD simulations data. These plots provide key evidence that inhibitor bound kinase domain coexisted with active and inactive states in the MD simulations timescales. The pharmacophore functional features of virtual screened hit molecules can arrest upregulated Axl kinase activity in cancerous cells. Based upon ADME data all the proposed pharmacophore-based screened hit molecules are qualify the pharmacodynamics and cell toxicity level in normal physiological process. The proposed hit molecules are synthetically viable and possess drug likeness for Axl kinase domain inhibition. Based upon docking and MD simulations data, the seven hit molecules bind to kinase active site to trigger the specific domain dynamics in Axl kinase. The upregulated Axl kinase can be arrested to influence specific regions such as activation loop, regulatory spine, α -helix, coordinated with the binding of hit molecules. The overexpressed Axl kinase domain is an effective target for cancer therapy that can be inhibited by the hit molecules in a site selective manner. The Lys567 – Glu585 residue salt bridge distance analysis identify the active and inactive states in Axl kinase domain during the 250ns MD simulations while bound to reference and hit molecules. The regulatory spine is broken as the hit molecules moved towards the shallow depths of kinase activation loop and α C-helix regions (regulatory site bound region). While the regulatory spine is retained for hit molecules that bind the catalytic site (hinge region).

6A



6B

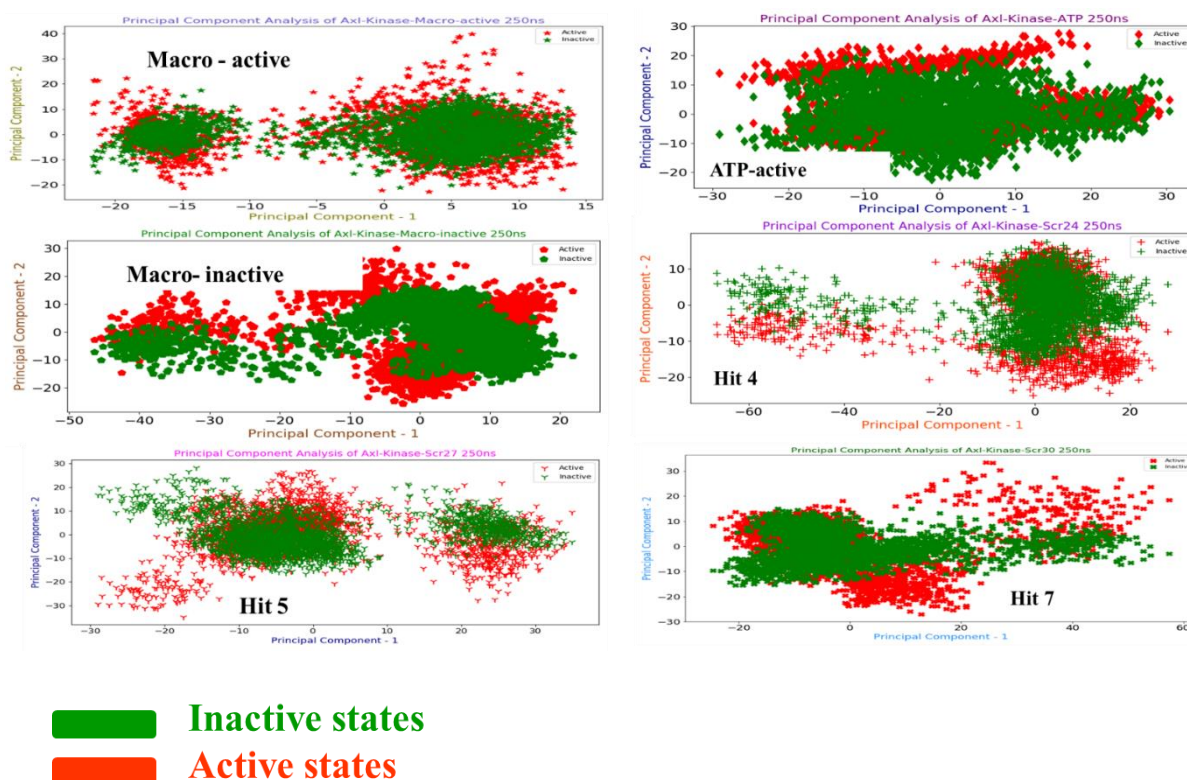


Figure 4.6A) Principal component analysis of Axl kinase regulatory site bound inhibitors
Figure 4.6B) Principal component analysis of Axl kinase hinge region bound inhibitors

4.3.7. Overall mechanism of kinase inhibition with screened small molecules

The Axl kinase domain structure is influenced by screened hit molecules and differentiated based upon their site selective affinity towards binding sites. The regulatory binding site (P-loop – α C helix - activation loop) is bound by specific Axl kinase inhibitors (hit 1, hit 2, hit 3, hit 6), while hit - 3 is converted into complete inactive state mode with broken regulatory spine. All regulatory site inhibitors coexist between the active – inactive states based on the distance between Lys567 – Glu585 that varies from 3.0 – 10.5 Å (hit 1, 2, 6 distance plots). The inactive states have high binding free energies among all active states (Table-4.3). The hinge region binding inhibitors bind similar to 7YS macrocyclic inhibitor in the 5U6B crystal structure. These hit molecules do not influence regulatory spine hydrophobic residues including 7YS. The distance plots describe the 7YS inhibitor influencing the active state of Axl kinase with ionic bond distance between Lys567 – Glu585 residues (4.5 Å) compare to inactive state (17.5 Å). In a similar manner, the hit 4, hit 7 existed as active states. However, hit 5 has sinusoid pattern in the ionic bond distance between Lys567 – Glu585 residues (ranging from 4.5 to 18 Å). These insights of Axl kinase domain is influenced by binding to different hit molecules in the presence water solvation system. The MM-GBSA and MM-PBSA data also support two different kinase inhibiting sites with seven screened hit molecules. This work provides a glimpse based on computational studies of Axl kinase domain inhibition by hit molecules designed from pharmacophore virtual screening using 7YS kinase inhibitor.

4.3.8. Non-bonding interactions of Axl RTKs with hinge region and regulatory site bound hit molecules

The amino acid residues that are in the vicinity of hit molecules in Axl RTK active site were examined. The interactions with corresponding hit molecules at active site differentiate with hinge region and regulatory site bound Axl kinase domain. The regulatory site bound hit molecules (hit 1, hit 2, hit 3 and hit 6) have interactions with Met589, Asp690, Phe691, Met679 residues. The hinge region bound hit molecules (Hit 4, hit 5 and Hit 7) interact with Met623, Gly626, Asp627 and Met679. The molecules hit 1, 2 and 6 hit molecules (active state) have no interaction with the hinge region residue Met623 of Axl kinase domain. But the inactive state hit 3 has interaction also with Met623. Similarly, the hit 5 interacts with the Lys567, Phe592, Met623, Asp690, Phe691, Gly692 and Ser694 in the Axl kinase domain. The uniqueness of hit molecules 3 and 5 is also observed from the hydrogen bond analysis. The inactive state of hit 3 and 5 have interacts with Met623 and Asp690 residues specifically despite their binding to regulatory or hinge region sites, respectively. These hydrogen bond analysis describe site specific inhibitor molecules to Axl kinase active site.

4.4. Conclusions

Protein kinases have diverse binding pockets based upon co-factor and substrate binding and allosteric binding. The pharmacophore features of macrocyclic inhibitor, 7YS taken as reference for virtual screening of chemical libraries, molecular docking and molecular dynamics simulations identified molecules that bind to regulatory and hinge region sites. The scaffolds with bicyclic and heterocyclic groups have made contribution to two different binding site regions in Axl kinase inactive state. The molecules hit 4, 5 and 7 bind at kinase hinge region which facilitate hydrogen bonding with the hinge region residues Met623, Gly626, Asp627 and Met679. The molecules hit 1, 2, 3 and 6 selectively bind to the kinase regulatory region and make hydrogen bonds with Met589, Met598, Asp690 and Met679 where as inactive state of hit 3 and hit 5 bind with Met623, Met679 and Asp690 in a concise manner. These hit molecules moved across from hinge region to regulatory region of shallow depth space between α C-helix and activation loop. But the hinge region bound kinase hit molecules have moved from kinase activation loop towards the hinge region where this space would be occupied by ATP co-factor in active state of Axl kinase. The hinge and regulatory bound Axl kinase active and inactive states are differentiated with regulatory spine mechanism. The (P-loop) Lys567-Glu585 (α C- helix) salt bridge distance analysis is akey to identify active and inactive states of Axl kinase. PCA describes the hidden states of Axl kinases with seven hit molecule and 7YS complexes. The systematic comparative analysis of hit molecules to mimic the arrest of overexpressed Axl kinases activity in various structural environments is studied computationally.

Chapter -5

Mutations in the receptor-binding domain of human SARS CoV-2 spike protein increases its affinity to bind human ACE-2 receptor

Abstract

The severe acute respiratory syndrome virus-2 (SARS CoV-2) infection is a disease causative agent of severe respiratory problems in humans and animals and it has resulted in the current global pandemic. The binding of SARS CoV-2 spike protein receptor-binding domain (RBD) to the human angiotensin converting enzyme-2 (ACE-2) receptor causes the host infection. The spike protein has undergone several mutations with reference to the initial strain isolated during December 2019 from Wuhan, China. A number of these mutant strains have been reported as variants of concern and as variants being monitored. Some of these mutants are known to be responsible for increased transmissibility of the virus. Here, the crystal structure of the RBD in complex with ACE-2 available in the public domain was used and analysed the 500 ns MD simulations of wild-type and mutants; G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493A, Q493R, G496S, Q498R, N501Y, Y505H. The ionic, hydrophobic and hydrogen bond interactions, residue flexibility, binding energies and structural variations are characterized. The MD simulations provide clues to the molecular mechanisms of ACE-2 receptor binding in wild-type and mutant complexes. The mutant spike protein RBD was associated with greater binding affinity with ACE-2 receptor.

5.1. Introduction

The severe acute respiratory syndrome coronavirus-2 (SARS CoV-2) has caused a global pandemic of the coronavirus disease 2019 (COVID-19) during the last 40 months. COVID-19 has been the single major cause of death due to any disease within a short span of time. SARS CoV-2 was first reported in individuals known to have been in contact with wildlife animals at the live animal and seafood market in Jiangnan District, Wuhan (Zhu et al., 2020). SARS CoV-2 is similar to SARS CoV (2003 to 2005), Middle East respiratory syndrome coronavirus (MERS CoV) (2012 to 2013) and other human CoVs in the 20th century that has led to epidemics resulting in severe respiratory diseases and deaths (Guruprasad, 2021a). These viruses harbour ~ 30K bp single stranded positive-sense RNA genome. SARS CoVs enter human cells through fusion of viral and host cellular membranes mediated by the interaction between viral spike protein and human angiotensin converting enzyme-2 (ACE-2) (Guruprasad, 2020; 2021b; Li et al., 2003; Shang et al., 2020). The SARS CoV-2 spike protein is a heavily glycosylated homo-trimeric protein with ~1,273 amino acids and the sequence region (amino acids 333-520) constitutes the receptor-binding domain (RBD) that interacts with human ACE-2 receptor. The three-dimensional structures of the spike protein apo and RBD bound forms to human ACE-2 receptor are available in the public domain (Wang et al., 2020; Xiao et al., 2021; Xu et al., 2021). Viruses acquire mutations over a period of time during host infection giving rise to new sequence variants. The mutations corresponding to the entire spike protein across the different lineages are shown in Table 5.1. The three-dimensional crystal structure of the spike protein RBD (residues 333-520) complexed with human ACE-2 is available in the PDB (Lan et al., 2020; Wang et al., 2020). The structure comprises a five-stranded antiparallel β -sheet. According to the PDBSum (Laskowski et al., 2018), the amino acids region 440 - 506 (67 amino acids) located between β 4 and β 5 strands folds into an extended loop that comprises short stretches of two α -helices (439-442 and 502-505) and anti-parallel β -sheets formed by two β -strands pairs (451-455 with 492-495) and (472-474 with 488-490). The secondary structural elements are connected by loops between 443-448, 476-488, 497-503. The binding between virus and host cell receptor are mediated via non-bonding interactions through RBMs in the RBD extended loop and the virus binding motifs (VBMs) on the ACE-2 receptor shown in Figure 5.1A. The substitution mutations in RBD were shown to increase the transmissibility of COVID-19 and decreased protection from vaccines (Bian et al., 2021; Chen et al., 2020; Gomez et al., 2021; Harvey et al., 2021; Noh et al., 2021; Zhou et al., 2021). This prompted analysis of the mutations and SIE estimations corresponding to variants

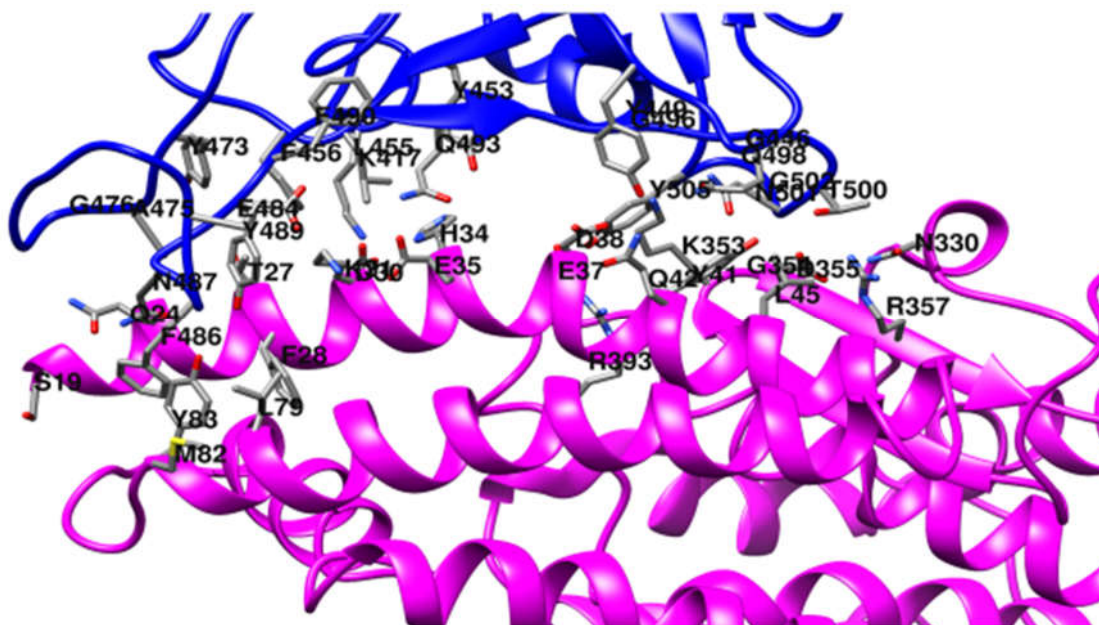


Figure 5.1A) Amino acid residues in 4.5 Å vicinity between human SARS CoV-2 spike protein RBD (blue) and human ACE-2 receptor (magenta). The hydrogen bonding interactions involved residues; Ala475-Ser19, Asn487-Gln24, Thr500-Tyr41, Lys417-Asp30, Tyr449-Asp38, Tyr449-Gln42, Asn487-Tyr83, Gln498-Gln42, Asn501-Tyr41, Gly502-Lys353

of concern and variants being monitored using MD simulations of the spike protein RBD domain and human ACE-2 receptor complex. These studies on host-virus protein-protein interactions at the atomic level suggest molecular mechanisms of their binding. Using molecular docking and MD simulations it has been shown that mutations in the SARS CoV-2 spike protein RBD are responsible for strong ACE-2 binding and poor anti-SARS CoV-2 monoclonal antibodies cross-neutralization (Shah et al., 2020). The alanine scanning mutagenesis and computational binding affinity studies of certain residues in SARS CoV-2 RBD complex with human ACE-2 showed that the mutations in conserved receptor binding motif (RBM) affects the structural-dynamics of the complex. The charge distribution disturbs the inter-molecular non-bonded contacts thereby perturbing the strength of binding to host cell ACE-2 receptor (Dehury et al., 2021). A pictorial representation of the non-bonding interactions between the human spike protein RBD and human ACE-2 described above are given in Table 5.2 are shown in Figure 5.2E.

The E484Q, L452R and double mutant E484Q and L452R were studied using MD simulations (Antony & Vijayan., 2021). The triple mutant variants; K417N□E484K□N501Y and K417T□E484K□N501Y studied using molecular docking and MD simulations attributed the increased binding of spike protein RBD to ACE-2 mainly due to the electrostatic contribution (Khan et

al., 2021). Single amino acid point mutations in spike protein RBD and C-terminus were studied using MD simulations (Ahamad, Kanipakam, & Gupta., 2020)

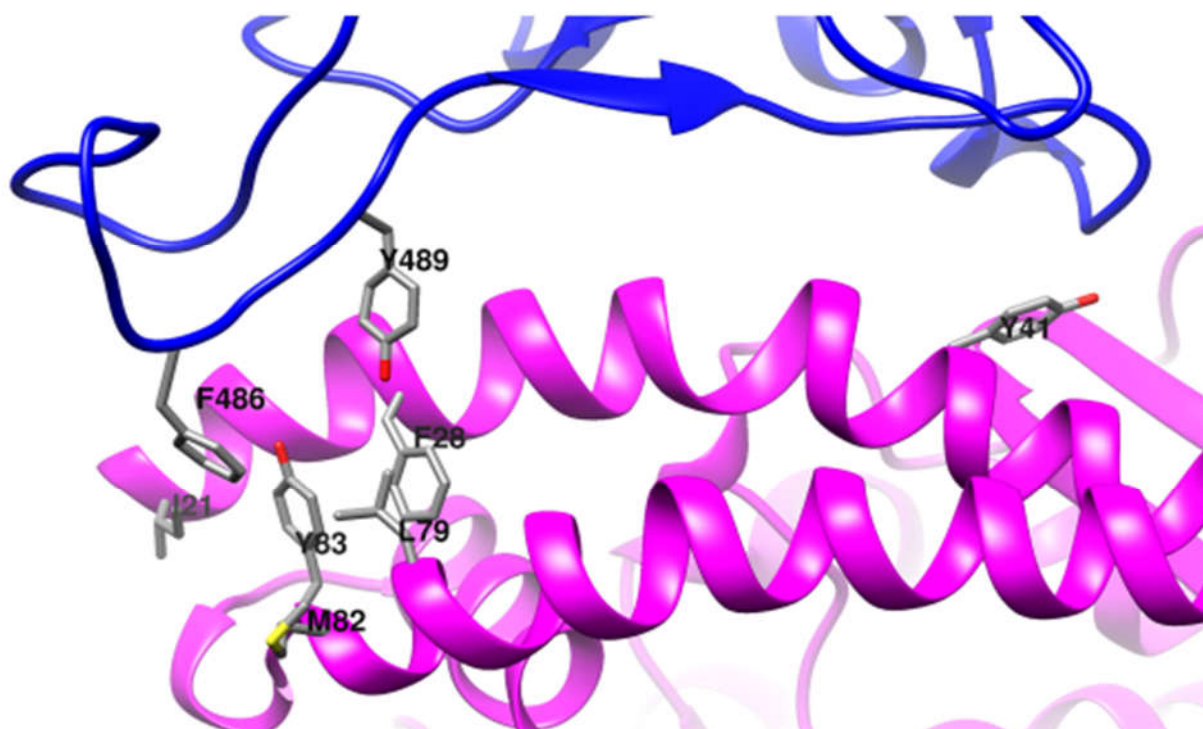


Figure 5.1B) Amino acid residues involved in hydrophobic interactions between human SARS CoV-2 Spike protein RBD (blue) and human ACE-2 receptor (magenta)

The substitution mutations in RBD were shown to increase the transmissibility of COVID-19 and decreased protection from vaccines (Harvey et al., 2021; Gomez, Perdiguero, & Esteban., 2021; Noh, Jeong, & Shin., 2021; Zhou et al., 2021; Bian et al., 2021; Chen et al., 2020). The mutations analysed with SIE (or the binding free energy) corresponding to variants of concern and variants being monitored using MD simulations of the spike protein RBD domain and human ACE-2 receptor complex. These studies on host-virus protein-protein interactions at the atomic level suggest molecular mechanisms of their binding.

5.2. Methods

5.2.1. Generation of mutants in human SARS CoV-2 spike protein RBD

The crystal structure of the human SARS CoV-2 spike protein RBD with human ACE-2 receptor complex (PDB_id: 6LZG) was used to generate single amino acid substitution mutations; K417N, N440K N501Y, L452R, T478K, E484K, S494P, K417A, Q498A, T500I using Discovery Studio 3.5.

5.2.2. Molecular dynamics simulations

The MD simulations of the wild-type and mutant RBD - ACE2 receptor heterodimer complexes were carried out using AMBER (ver.18.14) (Gotz et al., 2012; Salomon-Ferrer et al., 2013). The GAFF2 force fields charges and Amberff14sb protein atomic positions of all the systems were generated with Antechamber using am1bcc method (Lindorff-Larsen et al., 2010). All input parameter files for MD simulations were generated for the heterodimer by adding hydrogen atoms with H++ server (Anandakrishnan et al., 2012; Gordon et al., 2005). Sodium ions were added to the system in order to neutralize excessive charge generated by solvating the complex in a 10 Å cubic box. The final ionic concentration of the systems was set to 100 mM. The Amberff14sb-idln force fields was used with TIP3P water model (Jorgensen et al., 1983; Mark & Nilsson, 2001). The topology input.crd parameter files were generated with tLEaP module using the Amber suite (Hornak et al., 2006). The MD simulations were run at 300 K temperature and 1 atmospheric pressure. A 2 fs time step was considered for friction coefficient of 1/ps of Langevin integrator. Energy minimization was carried out using steepest descent method for 40,000 cycles in order to overcome short range null contacts in the system. The long-range electrostatic interactions were handled using the particle-mesh Ewald (PME) method with a 9 Å real-space cut-off and with PME order 4. The systems were double minimized under NPT ensemble at interval of 5 frames to maintain pressure of 1 atmosphere with Monte Carlo barostat. The minimization was carried out under NVT ensemble at interval of 5 frames, in order to maintain volume and temperature at 300 K using Monte Carlo thermostat (Darden et al., 1993; Wang et al., 2006). The systems were equilibrated for 7 ns before production run. The frames from production runs were saved every 10 ps. All molecular systems were executed for 250 ns MD simulations using AMBER.

5.2.3. Post MD data analysis

The MD simulations trajectory data analysis was carried out using cpptraj and pytraj in Amber tools. The average structures, RMSD and RMSF for all systems were derived from .trr analysis. The cpptraj h-bond sub-level trajectory analysis provides the average hydrogen bonding distance between human ACE-2 (chain-A) and spike protein RBD (chain-B). The SIE is an indirect method to calculate the binding free energies between protein-protein or protein-ligand complexes simulated with explicit solvent models. To understand the protein-protein affinities between wild-type and mutant proteins, the SIE-traj analysis (Naïm et al., 2007) was performed between ACE-2 receptor and RBD that provide insights into dynamical state of the non-bonding interactions. The entire 25,000 frames generated from 250 ns MD simulations was taken into consideration to calculate the SIE (Cui et al., 2008; Sulea & Purisima, 2012). The

SIEs and non-bonded interaction energies were computed for the wild-type and mutant dimeric RBD – ACE-2 complexes.

5.3. Results and Discussions

5.3.1. Molecular dynamics simulations

The MD simulations showed enhanced binding of human SARS CoV-2 spike protein RBD to human ACE-2 receptor, suggesting the possible role of mutations in leading the virus to become variants of concern or variants being monitored. The structural superposition of initial and final MD simulation structures showed relative displacement of B-chain (RBD) compared to A-chain (ACE-2), except for the protein with E484K mutation as shown in Figure 5.2A. The RMSD plots for representative structures shown in Figure 5.2B and for all the structures demonstrated that the structures stabilized within 25 ns of MD simulations. The RMSF plots for all structures shown in Figure 5.2C indicated regions of fluctuations in spike protein RBD and human ACE-2 complex. The region of high flexibility was observed between amino acid positions 475-487 that corresponds to RBM in the RBD of all complexes, except the protein with the S494P mutation that was observed to be relatively stable. Fluctuations were also observed in certain loop regions of human ACE-2 located away from RBD. The T478K and N501Y mutant proteins have largest fluctuations indicative of the dynamical structures. Some structures have fluctuations in loop regions corresponding to amino acid positions; 357-371, 382-388 and 423-430. The observations of large structural fluctuations in certain mutant proteins are explained later using cluster analyses from 250 ns AMBER MD simulations trajectories, shown in Table 5.2. The K417 residue is mutated to N417 in B.1.351 lineage variant of concern. This mutation is reported to contribute to loss of serum antibody neutralization (Collier et al., 2021). The K417 residue is located on α 3-helix between β 3 and β 4 strands and is close to the insertion loop in RBD. The side chain N ζ -atom of K417 makes ionic interactions with side chain atom of Asp30 located on α 1-helix in human ACE-2. The Lys417-Asp30 side chain ionic interaction observed in the wild-type and all the other mutants is lost owing to the K417N mutation. The N501Y mutation associated with lineages; B.1.1.7, B.1.351, P.1 binds human ACE-2 receptor with increased binding affinity (Luan et al., 2021). The N501 is located in loop connecting the α -helix preceding β 5-strand and its side chain amide nitrogen makes intermolecular hydrogen bonding interactions with side chain OH atom of Tyr41 in ACE-2.

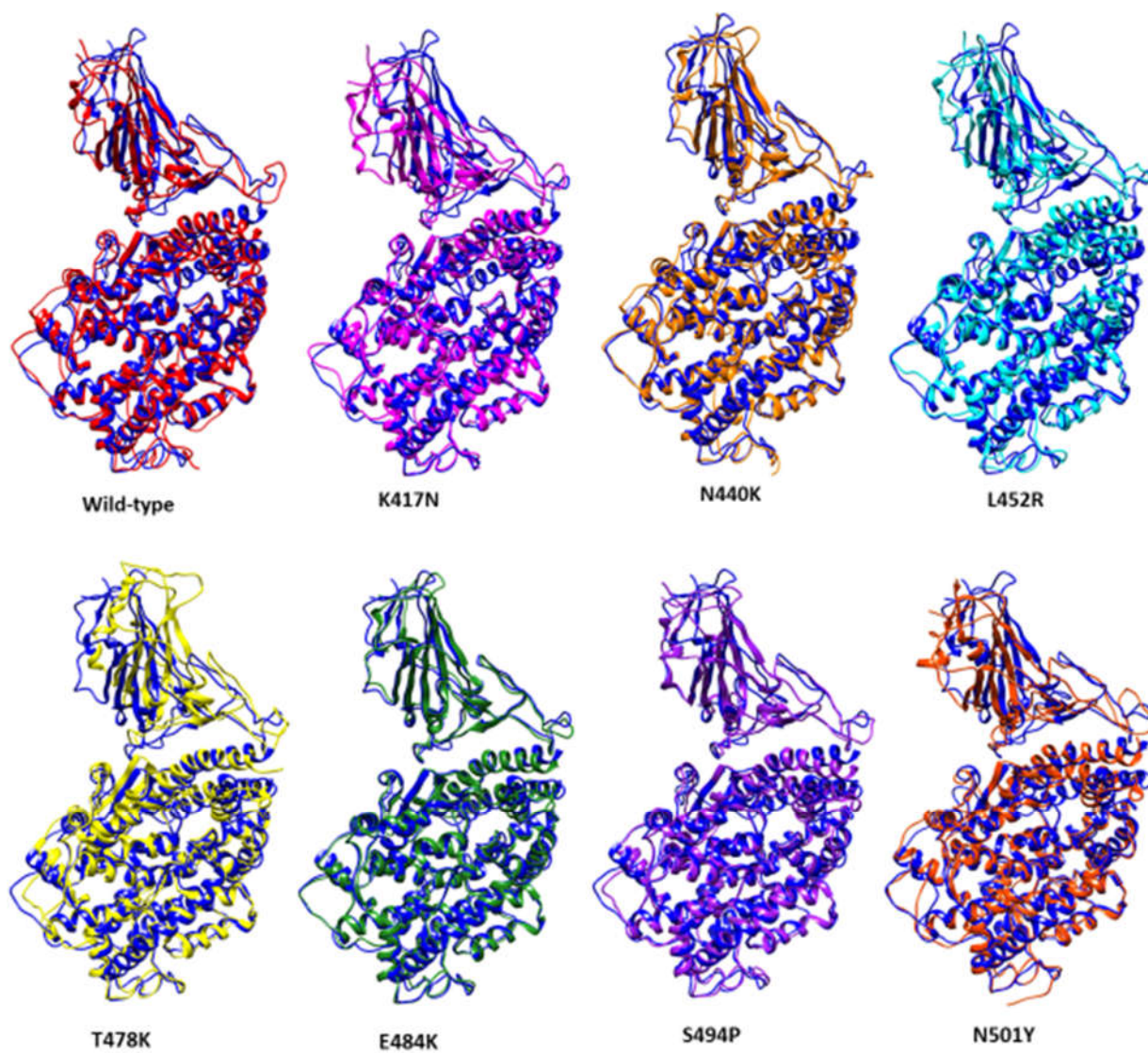
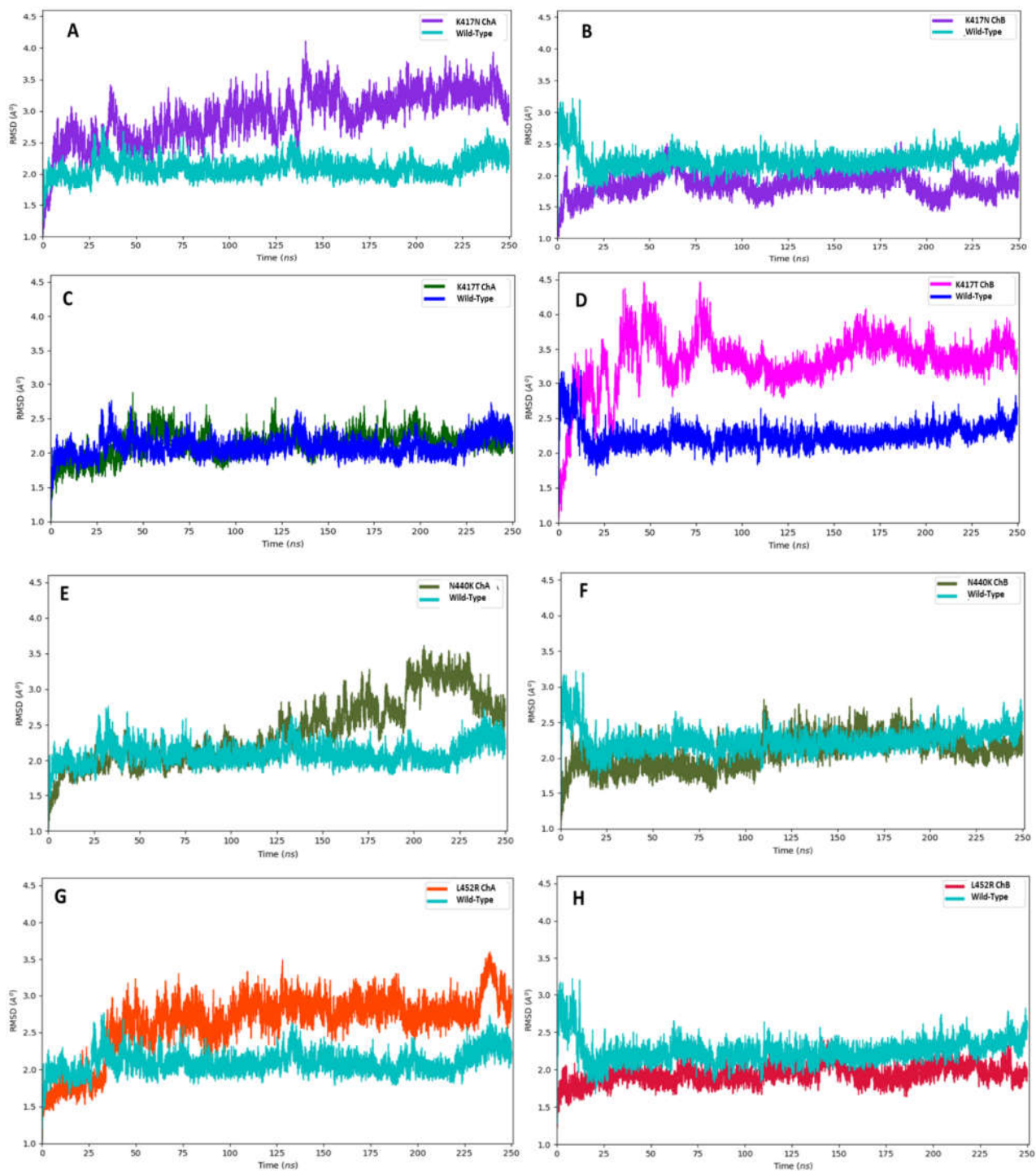


Figure 5.2A) Superposition of the initial structure (blue) with average MD simulations structure (Shown in different colours for the different mutations) in wild-type and mutant human SARS CoV-2 spike protein RBD – human ACE-2 receptor complexes.



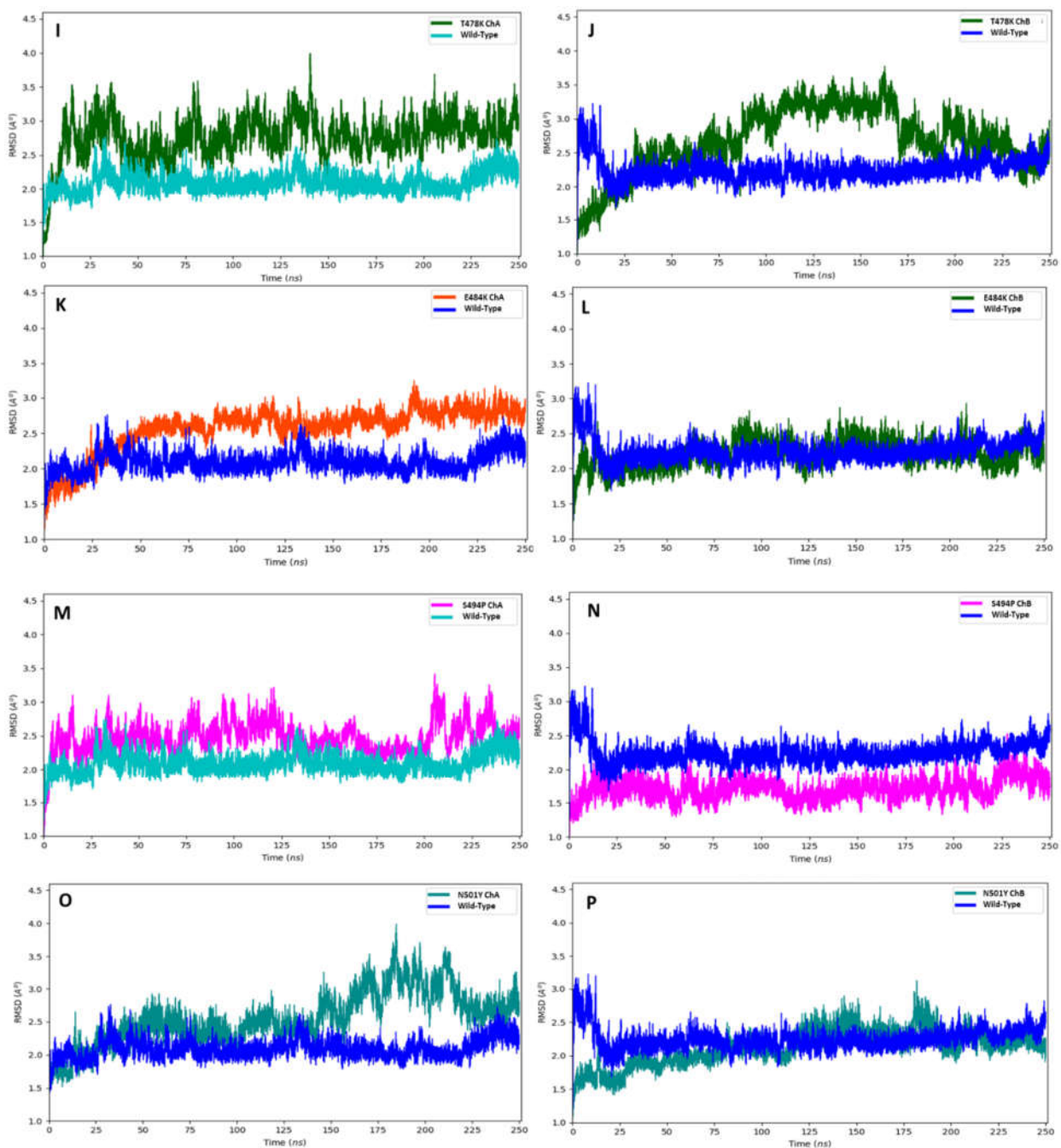


Figure 5.2B) The RMSD plots for wild-type and representative mutant human SARS CoV-2 spike Protein RBD chain-B (ChB) in complex with human ACE-2 receptor chain-A (ChA) for 250 ns MD simulations (A-P Hetero dimer)

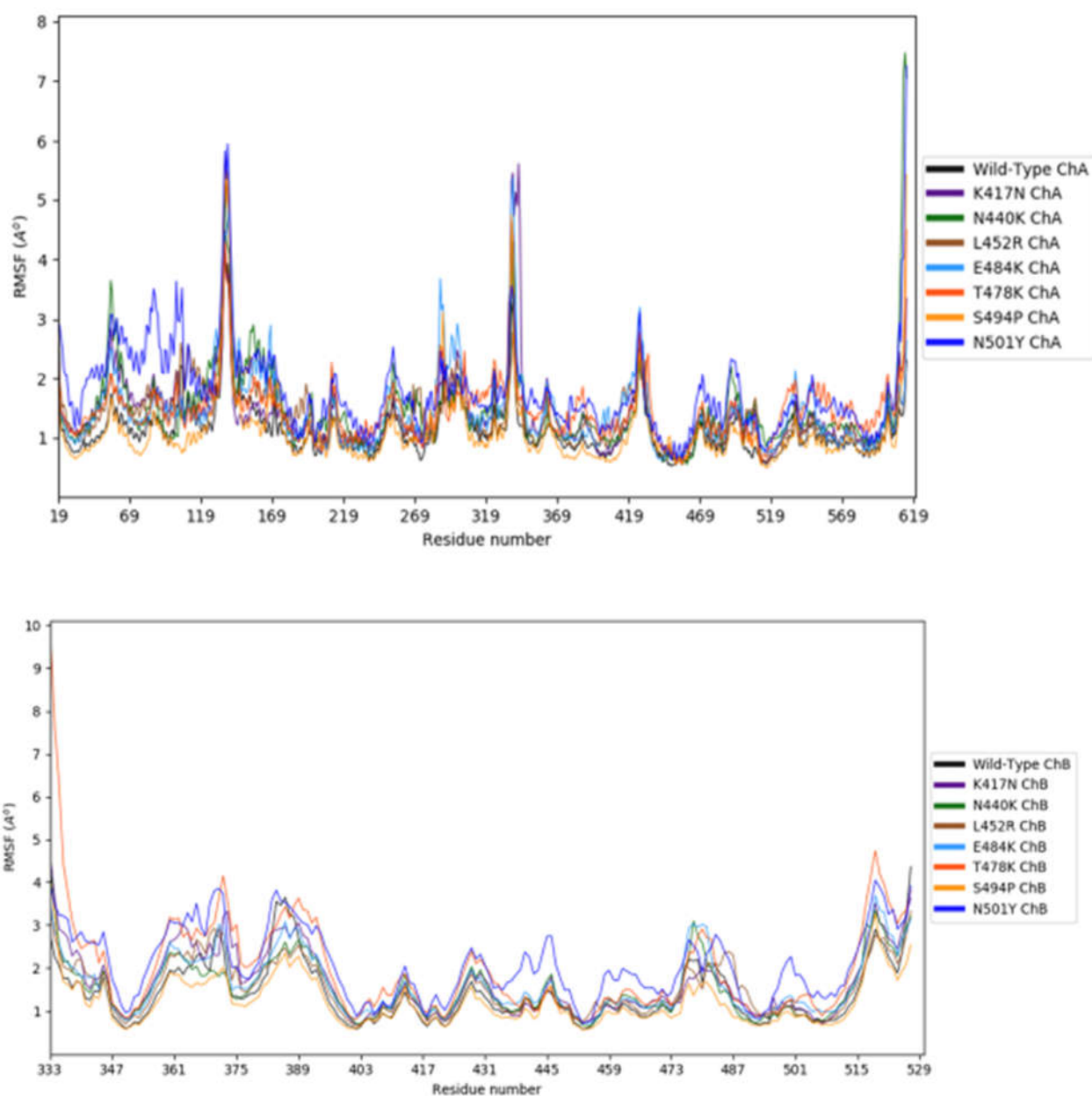


Figure 5.2C) The RMSF plots for wild-type and mutant SARS CoV-2 spike protein RBD (ChB) in complex with human ACE-2 receptor chain-A (ChA) for 250 ns MD simulations

5.3.2. Ionic, hydrophobic, hydrogen bond interactions in wild-type and mutant complexes

Several intermolecular interactions were observed to be mediated between the SARS CoV-2 spike protein RBD and human ACE-2 receptor in the wild-type and mutant protein complexes during the 250 ns MD simulations. Accordingly, these were classified as the hydrogen bond interactions, ionic interactions and the hydrophobic interactions. The SARS CoV-2 spike protein three-dimensional structure of the UK variant (B.1.1.7) in complex with ACE-2 heterodimeric domain showed π - π interaction between Tyr501 mutant and Tyr41 in ACE-2 that enhances the binding of spike protein to the receptor and abolishes binding of a potent neutralizing antibody (Yang et al., 2021). We observed that the N501Y mutant leads to hydrophobic interactions in mutant complex involving Tyr501-Tyr41 (Tyr41 is indicated in Figure 5.1B) at the protein-protein interface, in addition to the Phe486-Met82 hydrophobic interaction. The L452 is associated with RBM's anti-parallel β -sheet and is exposed to solvent and is not directly involved in interaction with ACE-2 (Luan et al., 2021). The mutation; L452R is observed in the lineages; B.1.427 and B.1.429. The T478 is associated with a loop in RBM. The hydrophobic cluster formed by Phe486 in RBD and its interactions with residues; Leu79, Met82, Tyr83 in ACE-2 is observed in the complexes with the mutants; T478K, E484K, N440K. The E484 is within a loop in RBM and does not interact with human ACE-2 in wild-type structure. The E484K mutant results in loss of serum antibody neutralization. Similar to E484, the S494 residue is also exposed to the solvent and is not involved in interactions with ACE-2. The S494 residue is associated with an antiparallel β -sheet in RBM. The L452 and S494 are located on the individual strands of antiparallel β -sheet and the N440 is located on a helical turn at the end of β 4-strand distant from the ACE-2 binding site.

In summary, Phe486 and Tyr489 in the human SARS CoV-2 spike protein were observed to be involved in hydrophobic interactions with residues; Ile21, Phe28, Tyr41, Leu79, Met82 and Tyr83 present on helices (H1, H2) of human ACE-2 as shown in Figure 5.1B. A list of the more common intermolecular hydrogen bond interactions for the wild-type and mutant SARS CoV-2 spike protein RBD – ACE-2 complex is provided in Table 5.1. The predominant hydrogen bond interactions were observed between; Asn487-Tyr83, Asn487-Gln24, Gln498-Gln24, Gln498-Gln42, Gln498-Tyr41, Gln498-Lys353, Gly496-Lys353, Gln493-Lys31, Ala475-Ser19 and Ala475-Gln24. The interactions involving the same amino acid residue with different partner residues in the complex during the MD simulations, suggests the promiscuous nature of hydrogen bonds in the complex of human SARS CoV-2 spike protein RBD and human ACE-2 receptor reflecting the structural plasticity associated with the RBM. Several hydrogen bonds, hydrophobic and ionic interactions mediate the intermolecular interactions

between the SARS CoV-2 spike protein RBD and human ACE-2 in wild-type and mutant proteins. Further, despite the mutations observed, the spike protein RBD is capable of interacting with the ACE-2 receptor leading to the host infection. Computational studies using MD simulations revealed additional π - π and π -cation interactions in the RBD-ACE-2 complexes (Tian et al., 2021). We observed that the N501Y mutant leads to hydrophobic interactions in the mutant complex involving Tyr501-Tyr41 (Tyr41 is indicated in Figure 5.1B) at the protein-protein interface, in addition to the Phe486-Met82 hydrophobic interaction. The L452 is associated with RBM's anti-parallel β - sheet, it is exposed to solvent and is not directly involved in interaction with ACE-2. The Leu452 together with Phe490 and Leu492 forms a hydrophobic surface on RBD. The E484Q, L452R and double mutant E484Q and L452R were studied using MD simulations (Antony & Vijayan, 2021). It was observed that Arg452 interacts more with neighboring residues Ser349, Tyr351, Phe490, Leu492 and Ser494 when compared to the wild-type and propose that the increased intra-molecular interactions could lead to the increased stability of the SARS CoV-2 spike protein. The mutation; L452R is observed in lineages; B.1.427 and B.1.429. The T478 is associated with a loop in RBM. The hydrophobic cluster formed by Phe486 in RBD and its interactions with residues; Leu79, Met82, Tyr83 in ACE-2 is observed in the complexes with the mutants; T478K, E484K, N440K. The E484 is within a loop in RBM and does not interact with human ACE-2 in wild-type structure. The E484K mutant results in the loss of serum antibody neutralization (Chen et al., 2021). Wang et al., 2021 have shown that the E484K mutation resulted in more favorable electrostatic interactions and significantly improved binding affinity with ACE-2. Further, the E484K mutation is shown to cause conformational rearrangements of the loop region containing the mutant residue that leads to a tighter binding with ACE-2 and formation of some new hydrogen bonds (Wang et al., 2021). Similar to E484, the S494 residue is also exposed to the solvent and is not involved in interactions with ACE-2. The S494 residue is associated with an antiparallel β -sheet in RBM. The L452 and S494 are located on the individual strands of antiparallel β -sheet and the N440 is located on a helical turn at the end of β 4-strand distant from the ACE-2 binding site. Two single amino acid substitution mutations (E484K, N501Y) and a triple mutant (K417N + E484K + N501Y) in the RBD domain in complex with human ACE-2 was studied using protein-protein docking and MD simulations (Istifli et al., 2021). The South African (K417N-E484K-N501Y) and Brazilian (K417T-E484K-N501Y) triple mutants have been shown to be lethal due to the inter-protein contacts specifically mediated via the electrostatic interactions from the results of molecular docking and MD simulations studies (Istifli et al., 2021; Khan et al., 2021). Single amino acid point mutations in the RBD and C-

terminus of spike protein were studied using MD simulations (Ahamad et al., 2022;; Istifli et al., 2021). These authors have shown that mutation brings about higher fluctuations mainly in the spike protein RBD region around 400-544 and heptad repeat 1 around 930-940 (Ahamad et al., 2022). Mutations in the SARS CoV-2 spike protein RBD are responsible for strong ACE-2 binding and poor anti-SARS CoV-2 monoclonal antibodies cross-neutralization (Shah et al., 2020). The alanine scanning mutagenesis and computational binding affinity studies of certain residues in SARS CoV-2 RBD complex with human ACE-2 showed that the mutations in conserved receptor binding motif (RBM) affects the structural dynamics of the complex. The charge distribution disturbs the inter-molecular non-bonded contacts thereby perturbing the strength of binding to host cell ACE-2 receptor (Dehury et al., 2021). A pictorial representation of the non-bonding interactions between the human spike protein RBD and human ACE-2 described above and in Table 5.2 are shown in Figure 5.1A. In summary, Phe486 and Tyr489 in the human SARS CoV-2 spike protein were observed to be involved in hydrophobic interactions with residues; Ile21, Phe28, Tyr41, Leu79, Met82 and Tyr83 present on helices (H1, H2) of human ACE-2 as shown in Figure 5.1B. A list of the more common inter-molecular hydrogen bonding interactions for the wild-type and mutant SARS CoV-2 spike protein RBD – ACE-2 complex is provided in Table 5.1. The predominant hydrogen bonding interactions were observed between; Asn487-Tyr83, Asn487-Gln24, Gln498-Gln24, Gln498-Gln42, Gln498-Tyr41, Gln498-Lys353, Gly496-Lys353, Gln493-Lys31, Ala475-Ser19 and Ala475-Gln24. The interactions involving the same amino acid residue with different partner residues in the complex during the MD simulations, suggests the promiscuous nature of hydrogen bonds in the complex of human SARS CoV-2 spike protein RBD and human ACE-2 receptor reflecting the structural plasticity associated with the RBM. Several hydrogen bonds, hydrophobic and ionic interactions mediate the intermolecular interactions between the SARS CoV-2 spike protein RBD and human ACE-2 in wild-type and mutant complexes. Further, despite the mutations observed, the spike protein RBD is capable of interacting with the ACE-2 receptor leading to the host infection. Each of the single amino acid substitution mutations independently have significant effect on the nature of interactions with human ACE-2. It has already been reported that single and double mutants in the RBD do not disrupt the interactions with ACE-2, but reduce the binding free energies because of the multiple interactions in the inter-molecular interactions and the extended molecular surface (Taka et al., 2021).

5.3.3. Cluster analysis

The flexible partners stabilizing the intermolecular interactions between the heterodimeric complexes were further analyzed using cluster analyses shown in Figure 5.2D. The population of clusters and their standard deviations is provided in the Table 5.2. This table indicates 6 clusters with greater than 0.05 fraction (1,250 populations in a given cluster out of the 25,000 frames) in wild-type protein, whereas, the clusters associated with the different mutations were; K417N (8 clusters), T478K (8), S494P (10), N440K (10), L452R (9), N501Y (7), E484K (9). The members of each cluster comprise structurally similar conformations. The presence of only 6 clusters in the wild-type protein with a higher population of frames suggests the structural stability of wild-type complex, which contrasts with the relatively higher dynamics observed for the mutant proteins associated with larger number of clusters.

5.3.4. Solvated interaction energies

To quantify the strength of intermolecular interactions the binding free energies of the heterodimers were analyzed. The SIE binding free energy calculations were carried out on the AMBER MD simulation trajectories shown in Table 5.3. Among all complexes studied, low SIE values were observed for S494P (-31.24 kcal/mol), T478K (-29.67 kcal/mol), K417N (-29.59 kcal/mol), L452R mutant (-27.94 kcal/mol), N440K (-20.18 kcal/mol), E484K (-19.15 kcal/mol) and N501Y (-18.98 kcal/mol) compared to the binding free energy for the wild-type heterodimeric complex (Zhang et al., 2021) that was -13.75 kcal/mol. It was observed that all mutations in the spike protein RBD were associated with lower binding free energies compared to the wild-type proteins indicating better binding efficiency to human ACE-2. All mutations attributed as variants of concern or variants being monitored in RBD are known to increase transmissibility. These mutations cause greater infectivity to the host and may be under positive selection pressure. However, for some mutations, such as, K417A, Q498A, T500I in spike protein RBD previously reported (Guruprasad, 2021b), relatively higher binding free energies compared to the wild-type hetero-dimeric complex were observed (Table 5.3). These mutant proteins with implied reduced binding affinity to human ACE-2 may therefore not be significant mutations. The SIE binding free energy (ΔG) and its components; van der Waals interaction energy (vdW); Coulomb interaction energy (Coul); Reaction Field (RF); Constant (Const). Thereby, mutations resulting in greater infectivity to host seem to have been selected in the evolution of human SARS CoV-2 spike protein. Such mutations have therefore become prominent and have resulted as variants of concern or variants being monitored.

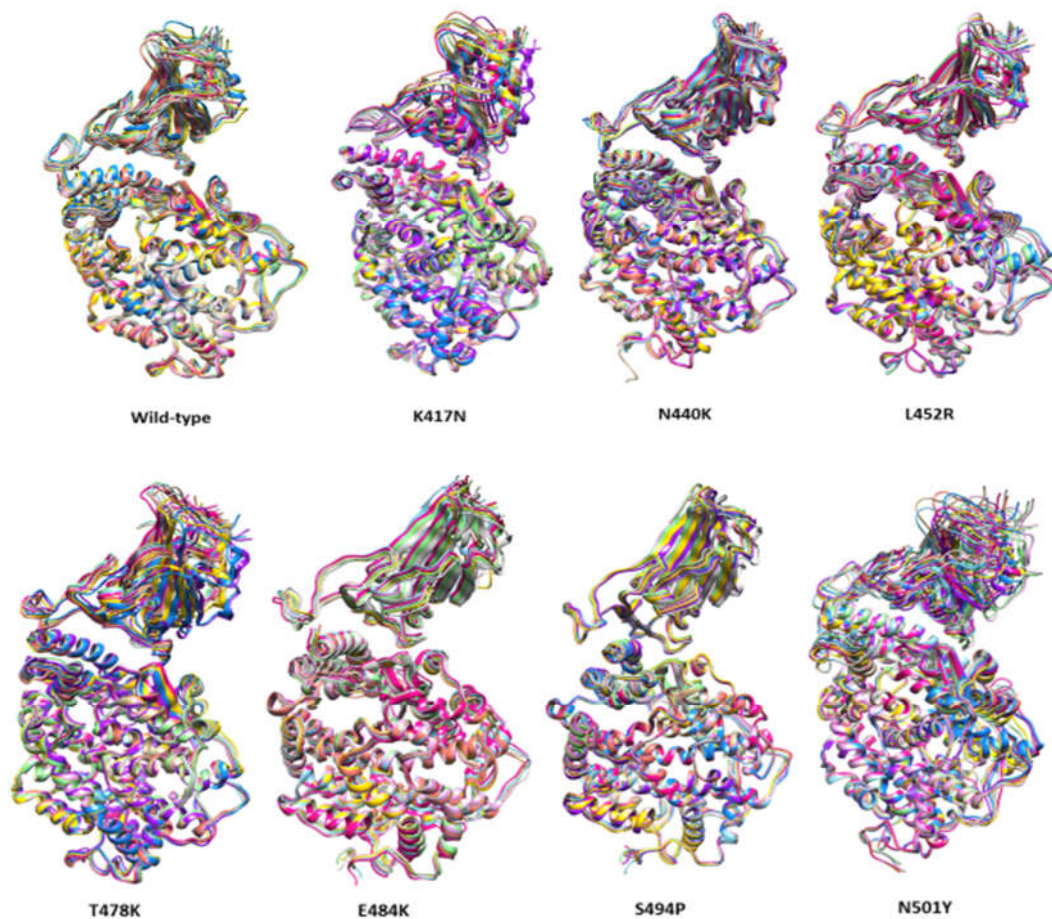


Figure 5.2D) Cluster analyses for wild-type and mutant SARS CoV-2 spike protein RBD and Human ACE-2 receptor for 250 ns MD simulations data.

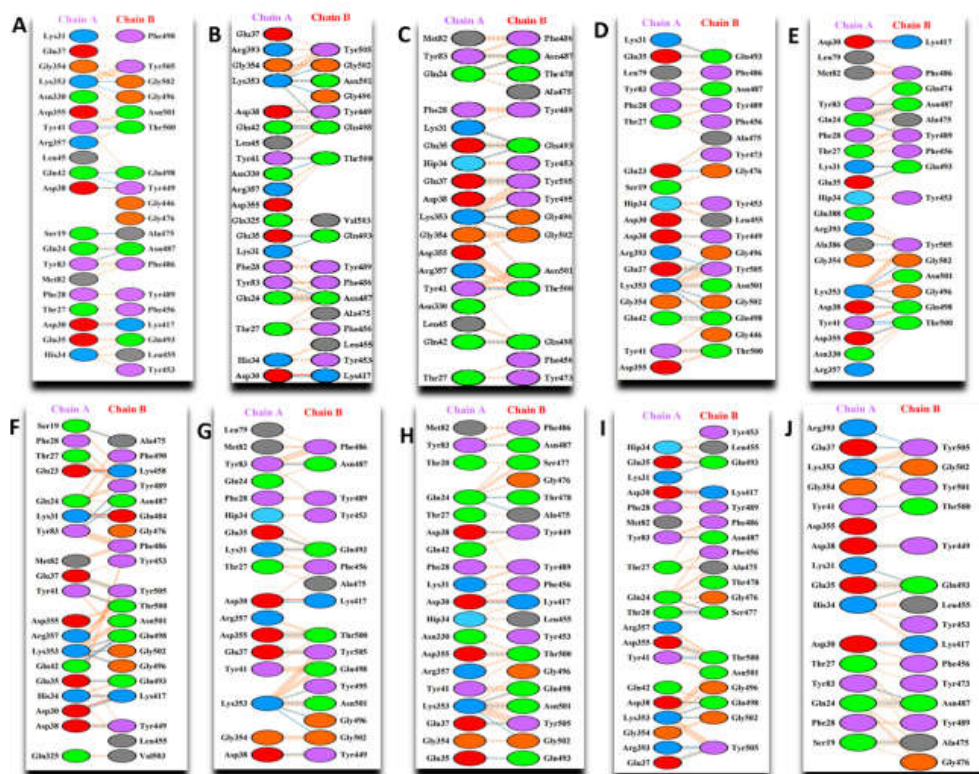


Figure 5.2E) Representation of non-bonding interactions between the human ACE-2 and spike protein RBD in the initial, wild-type and mutant complexes after 250 ns MD simulations. (initial structure PDB id: 6LZG, chains-A and B), B- wild-type, C- K417N, D-K417T, E-N440K, F-L452R, G-T478K, H-E484K, I-S494P, J-N501Y;

Lineage Variants being monitored	WHO label	Mutations relative to reference protein in human SARS CoV-2 spike protein RBD
B.1.1.7	Alpha	69del, 70del, 144del, E484K, S494P, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H, K1191N
B.1.351	Beta	D80A, D215G, 241del, 242del, 243del, K417N, E484K, N501Y, D614G, A701V
P.1	Gamma	L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I
B.1.427 B.1.429	Epsilon	L452R, D614G S13I, W152C, L452R, D614G
B.1.525	Eta	A67V, 69del, 70del, 144del, E484K, D614G, Q677H, F888L
B.1.526	Iota	L5F, D80G, T95I, 144del, F157S, D253G, L452R, S477N, E484K, D614G, A701V, T859N, D950H, Q957R
B.1.617.1	Kappa	T95I, G142D, E154K, L452R, E484Q, D614G, P681R, Q1071H
B.1.617.3	Not applicable	L452R, E484Q, D614G, P681R
P.2/P.1.1.28.2	Zeta	E484K, D614G, T859I, Y1176F
B.1.621	Mu	R346K, E484K, N501Y, D614G, P681H
<u>Variant of concern</u>		
B.1.617.2	Delta	T19R, V70F, T95I, G142D, 156-157del, R158G, A222V, W258L, K417N, L452R, T478K, D614G, P681R, D950N
B.1.1.529	Omicron	A67V, Δ69-70, T95I, G142D, Δ143-145, Δ211, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N N440K G446S S477N T478K F484A

Table 5.1) The deletions and substitution mutations in human SARS CoV-2 spike protein according to the different lineages.

Protein	Non-bonding interactions; residues involved in RBD-ACE-2
Wild-type	Hydrogen bonds: Asn487-Gln24, Gln493-Lys31, Gln493-Glu35, Tyr449-Asp38, Gln498-Gln42, Thr500-Tyr41, Asn487-Tyr83, Asn501-Lys353, Gly496-Lys353, Gln498-Lys353, Asn501-Lys353, Gly502-Lys353. Ionic bond: Lys417-Asp30. Hydrophobic: Tyr489-Phe28, Tyr489-Leu79, Tyr489-Tyr83.
K417N	Hydrogen bonds: Asn487-Gln24, Gln493-Lys31, Gln493-His34, Gln493-Glu35, Tyr505-Glu37, Asn501-Tyr41, Thr500-Tyr41, Gln498-Gln42, Asn487-Tyr83, Tyr495-Lys353, Gly496-Lys353, Gly502-Lys353, Thr500-Asp355. Hydrophobic: Tyr489-Phe28, Phe486-Met82 and Phe486-Tyr83.
N501Y	Hydrogen bonds: Ala475-Ser19, Asn487-Gln24, Gln493-Lys31, Gln493-His34, Gln493-Glu35, Tyr505-Glu37, Tyr449-Asp38, Thr500-Tyr41, Asn487-Tyr83, Gly502-Lys353, Tyr505-Arg393. Ionic: Lys417-Asp30. Hydrophobic: Tyr501-Tyr41, Phe486-Met82.
L452R	Hydrogen bonds: Ala475-Ser19, Tyr489-Gln24, Tyr453-His34, Lys417-His34, Gln493-Glu35, Tyr505-Glu37, Tyr449-Asp38, Thr500-Tyr41, Gln498-Gln42, Asn487-Tyr83, Tyr489-Tyr83, Gly496-Lys353, Gln498-Lys353, Gly502-Lys353. Ionic: Lys417-Asp30, Lys458-Glu23, Glu484-Lys31. Hydrophobic: Phe486-Ile21, Phe486-Tyr83, Tyr489-Phe28.
T478K	Hydrogen bonds: Gln493-Lys31, Gln493-Glu35, Tyr505-Glu37, Tyr449-Asp38, Asn487-Tyr83, Tyr495-Lys353, Gly496-Lys353, Asn501-Lys353, Thr500-Asp355, Gly502-Lys353. Ionic: Lys417-Asp30. Hydrophobic: Phe486-Leu79, Phe486-Met82, Phe486-Tyr83.
E484K	Hydrogen bonds: Ala475-Gln24, Thr478-Gln24, Gln493-Glu35, Tyr505-Glu37, Tyr449-Asp38, Asn487-Tyr83, Gln498-Lys353, Asn501-Lys353, Gly502-Lys353, Thr500-Asp355. Ionic: Lys417-Asp30. Hydrophobic: Phe486-Leu79, Phe486-Met82, Phe486-Tyr83.
S494P	Hydrogen bonds: Ser477-Thr20, Ser477-Gln24, Gln493-Lys31, Gln493-Glu35, Tyr505-Glu37, Gln498-Asp38, Thr500-Tyr41, Asn487-Tyr83, Gln498-Lys353, Gln498-Lys353, Gly502-Lys353, Thr500-Asp355, Thr500-Arg357, Tyr505-Arg393. Ionic: Lys417-Asp30. Hydrophobic: Phe486-Met82, Phe486-Tyr83.
N440K	Hydrogen bonds: Asn487-Gln24, Tyr489-Gln24, Ala475-Thr27, Gln493-Lys31, Gln493-Glu35, Gln498-Asp38, Thr500-Tyr41, Asn487-Tyr83, Gly496-Lys353, Gln498-Lys353, Gly502-Lys353, Thr500-Asp355, Tyr505-Ala386, Tyr505-Arg393. Ionic: Lys417-Asp30. Hydrophobic: Phe486-Leu79, Phe486-Met82, Phe486-Tyr83.

Table 5.2) The non-bonding intermolecular interactions between human SARS CoV-2 spike protein RBD and human ACE-2 complex.

S. No	Mutation	vdW (kcal/mol)	Std Err	Coul kcal/mol)	Std Err	RF kcal/mol)	Std Err	Cavity kcal/mol)	Std Err	Const	ΔG kcal/mol)	Std Err
1	Wild Type	-92.19	0.15	-321.87	0.41	327.05	0.38	-16.67	0.02	-2.89	-13.75	0.02
2	K417N	-95.94	0.12	-177.17	0.23	34.54	0.28	-16.33	0.02	-2.89	-29.59	0.02
3	N440K	+48.36	0.84	-281.75	0.29	233.27	0.27	-17.17	0.02	-2.89	-20.18	0.81
4	L452R	+71.35	1.77	-420.37	0.51	341.95	0.40	-17.63	0.02	-2.89	-27.94	1.68
5	T478K	-83.46	0.12	-308.00	0.31	150.45	0.32	-14.65	0.02	-2.89	-29.67	0.03
6	E484K	-87.37	0.12	-439.40	0.35	387.00	0.34	-15.49	0.02	-2.89	-19.15	0.02
7	S494P	-94.05	0.12	-296.72	0.24	136.60	0.32	-16.50	0.01	-2.89	-31.24	0.03
8	N501Y	-84.69	0.02	-343.51	0.03	289.05	0.55	-14.49	0.04	-2.89	-18.98	0.06
9	K417T	-90.00	0.12	-204.21	0.24	179.70	0.23	-16.17	0.01	-2.89	-16.58	0.02

Table 5.3) Solvent interaction energies (SIE) in kcal/mol calculated from sietraj for the 250 ns MD simulations trajectories for wild-type and mutant human SARS CoV-2 spike protein RBD – ACE-2 complexes.

5.4. Conclusions

The MD simulation studies of the human SARS CoV-2 spike protein RBD and ACE-2 receptor complex for wild-type and mutants; K417N, N440K, N501Y, L452R, T478K, E484K, S494P reveal the molecular interactions underlying their binding affinities. The promiscuous nature of the non-bonding interactions is facilitated by the structural plasticity of the RBD that is accompanied by large conformational changes during the MD simulations. The mutant proteins are characterized by larger number of clusters indicating greater conformational variability between the proteins in different clusters. This suggests that the mutant proteins undergo relatively greater conformational changes compared to the wild-type proteins. The SIE analyses of human SARS CoV-2 spike protein RBD and ACE-2 complex suggests the basis for positive selection of mutants that have led to more infectious variants resulting in rapid spread of the COVID-19 disease.

List of References

1. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M. In {TensorFlow}: a system for {Large-Scale} machine learning, 12th USENIX symposium on operating systems design and implementation (OSDI 16), 2016; pp 265-283.
2. Abdullah, M.; Guruprasad, L., Structural insights into the inhibitor binding and new inhibitor design to Polo-like kinase-1 Polo-box domain using computational studies. *J Biomol Struct Dyn* 2019, 37 (13), 3410-3421.
3. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J Mol Biol* 1990, 215 (3), 403-10.
4. Ahamad, S.; Kanipakam, H.; Gupta, D., Insights into the structural and dynamical changes of spike glycoprotein mutations associated with SARS-CoV-2 host receptor binding. *J Biomol Struct Dyn* 2022, 40 (1), 263-275.
5. Akalu, Y. T.; Rothlin, C. V.; Ghosh, S., TAM receptor tyrosine kinases as emerging targets of innate immune checkpoint blockade for cancer therapy. *Immunol Rev* 2017, 276 (1), 165-177.
6. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J Mol Biol* 1990, 215 (3), 403-10.
7. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V., H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res* 2012, 40 (Web Server issue), W537-41.
8. Anderson, A. C., The process of structure-based drug design. *Chem Biol* 2003, 10 (9), 787-97.
9. Andrychowicz, M.; Denil, M.; Colmenarejo, S. G.; Hoffman, M. W.; Pfau, D.; Schaul, T.; Shillingford, B.; de Freitas, N., Learning to learn by gradient descent by gradient descent. *Adv Neur In* 2016, 29, 3988–3996.
10. Antony, P.; Vijayan, R., Molecular Dynamics Simulation Study of the Interaction between Human Angiotensin Converting Enzyme 2 and Spike Protein Receptor Binding Domain of the SARS-CoV-2 B.1.617 Variant. *Biomolecules* 2021, 11 (8).
11. Asensio, J. L.; Martin-Pastor, M.; Jimenez-Barbero, J., The use of CVFF and CFF91 force fields in conformational analysis of carbohydrate molecules. Comparison with AMBER molecular mechanics and dynamics calculations for methyl alpha-lactoside. *Int J Biol Macromol* 1995, 17 (3-4), 137-48.
12. Bager, P.; Wohlfahrt, J.; Fonager, J.; Rasmussen, M.; Albertsen, M.; Michaelsen, T. Y.; Moller, C. H.; Ethelberg, S.; Legarth, R.; Button, M. S. F.; Gubbels, S.; Voldstedlund, M.; Molbak, K.; Skov, R. L.; Fomsgaard, A.; Krause, T. G.; Danish Covid-19 Genome, C., Risk

- of hospitalisation associated with infection with SARS-CoV-2 lineage B.1.1.7 in Denmark: an observational cohort study. *Lancet Infect Dis* 2021, 21 (11), 1507-1517.
13. Bakan, A.; Bahar, I., The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proc Natl Acad Sci U S A* 2009, 106 (34), 14349-54.
 14. Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; Bahar, I., Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* 2014, 30 (18), 2681-3.
 15. Bakan, A.; Meireles, L. M.; Bahar, I., ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* 2011, 27 (11), 1575-7.
 16. Balasubramanian, P. K.; Balupuri, A.; Bhujbal, S. P.; Cho, S. J., 3D-QSAR-aided design of potent c-Met inhibitors using molecular dynamics simulation and binding free energy calculation. *J Biomol Struct Dyn* 2019, 37 (8), 2165-2178.
 17. Bates, S. E.; Fojo, A. T.; Weinstein, J. N.; Myers, T. G.; Alvarez, M.; Pauli, K. D.; Chabner, B. A., Molecular targets in the National Cancer Institute drug screen. *J Cancer Res Clin Oncol* 1995, 121 (9-10), 495-500.
 18. Bellan, M.; Pirisi, M.; Sainaghi, P. P., The Gas6/TAM System and Multiple Sclerosis. *Int J Mol Sci* 2016, 17 (11), 1807.
 19. Bennett, C. H., Efficient Estimation of Free-Energy Differences from Monte-Carlo Data. *J Comput Phys* 1976, 22 (2), 245-268.
 20. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L., The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res* 2007, 35 (Database issue), D301-3.
 21. Bian, L.; Gao, F.; Zhang, J.; He, Q.; Mao, Q.; Xu, M.; Liang, Z., Effects of SARS-CoV-2 variants on vaccine efficacy and response strategies. *Expert Rev Vaccines* 2021, 20 (4), 365-373.
 22. Bittrich, S.; Rose, Y.; Segura, J.; Lowe, R.; Westbrook, J. D.; Duarte, J. M.; Burley, S. K., RCSB Protein Data Bank: improved annotation, search and visualization of membrane protein structures archived in the PDB. *Bioinformatics* 2022, 38 (5), 1452-1454.
 23. Bordoli, L.; Kiefer, F.; Arnold, K.; Benkert, P.; Battey, J.; Schwede, T., Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* 2009, 4 (1), 1-13.
 24. Bosurgi, L.; Bernink, J. H.; Delgado Cuevas, V.; Gagliani, N.; Joannas, L.; Schmid, E. T.; Booth, C. J.; Ghosh, S.; Rothlin, C. V., Paradoxical role of the proto-oncogene Axl and Mer receptor tyrosine kinases in colon cancer. *Proc Natl Acad Sci U S A* 2013, 110 (32), 13091-6.

25. Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E., Scalable algorithms for molecular dynamics simulations on commodity clusters. In *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, Association for Computing Machinery: Tampa, Florida, 2006; pp 84–es.
26. Bowie, J. U.; Luthy, R.; Eisenberg, D., A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 1991, 253 (5016), 164-70.
27. Bowman, G. R.; Pande, V. S., Protein folded states are kinetic hubs. *Proc Natl Acad Sci U S A* 2010, 107 (24), 10890-5.
28. Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: the biomolecular simulation program. *J Comput Chem* 2009, 30 (10), 1545-614.
29. Broughton, M.; Verdon, G.; McCourt, T.; Martinez, A. J.; Yoo, J. H.; Isakov, S. V.; Massey, P.; Halavati, R.; Niu, M. Y.; Zlokapa, A., Tensorflow quantum: A software framework for quantum machine learning. *arXiv preprint arXiv:2003.02989* 2020.
30. Burley, S. K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G. V.; Duarte, J. M.; Dutta, S.; Fayazi, M.; Feng, Z.; Flatt, J. W.; Ganesan, S. J.; Goodsell, D. S.; Ghosh, S.; Kramer Green, R.; Guranovic, V.; Henry, J.; Hudson, B. P.; Lawson, C. L.; Liang, Y.; Lowe, R.; Peisach, E.; Persikova, I.; Piehl, D. W.; Rose, Y.; Sali, A.; Segura, J.; Sekharan, M.; Shao, C.; Vallat, B.; Voigt, M.; Westbrook, J. D.; Whetstone, S.; Young, J. Y.; Zardecki, C., RCSB Protein Data Bank: Celebrating 50 years of the PDB with new tools for understanding and visualizing biological macromolecules in 3D. *Protein Sci* 2022, 31 (1), 187-208.
31. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* 2007, 126 (1), 014101.
32. CA, V. O. N. K.; Merseburger, A. S.; Kuczyk, M. A., Novel therapeutic options for second-line therapy in metastatic renal cell carcinoma. *Mol Clin Oncol* 2016, 4 (6), 903-908.
33. Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; Fadda, E.; Amaro, R. E., Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent Sci* 2020, 6 (10), 1722-1734.
34. Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *J Comput Chem* 2005, 26 (16), 1668-88.

35. Castro, A.; Carter, H.; Zanetti, M., Potential global impact of the N501Y mutation on MHC-II presentation and immune escape. *bioRxiv* 2021, 2021.02.02.429431.
36. Cavasotto, C. N.; Phatak, S. S., Homology modeling in drug discovery: current trends and applications. *Drug Discov Today* 2009, 14 (13-14), 676-83.
37. Chatzou, M.; Magis, C.; Chang, J. M.; Kemena, C.; Bussotti, G.; Erb, I.; Notredame, C., Multiple sequence alignment modeling: methods and applications. *Brief Bioinform* 2016, 17 (6), 1009-1023.
38. Chen, H.; Cheng, F.; Li, J., iDrug: Integration of drug repositioning and drug-target prediction via cross-network embedding. *PLoS Comput Biol* 2020, 16 (7), e1008040.
39. Chen, J.; Wang, J.; Zhu, W., Molecular Mechanism and Energy Basis of Conformational Diversity of Antibody SPE7 Revealed by Molecular Dynamics Simulation and Principal Component Analysis. *Sci Rep* 2016, 6 (1), 36900.
40. Chen, J.; Wang, R.; Wang, M.; Wei, G. W., Mutations Strengthened SARS-CoV-2 Infectivity. *J Mol Biol* 2020, 432 (19), 5212-5226.
41. Chen, R.; Li, L.; Weng, Z., ZDOCK: an initial-stage protein-docking algorithm. *Proteins* 2003, 52 (1), 80-7.
42. Chen, R. E.; Zhang, X.; Case, J. B.; Winkler, E. S.; Liu, Y.; VanBlargan, L. A.; Liu, J.; Errico, J. M.; Xie, X.; Suryadevara, N.; Gilchuk, P.; Zost, S. J.; Tahan, S.; Droit, L.; Turner, J. S.; Kim, W.; Schmitz, A. J.; Thapa, M.; Wang, D.; Boon, A. C. M.; Presti, R. M.; O'Halloran, J. A.; Kim, A. H. J.; Deepak, P.; Pinto, D.; Fremont, D. H.; Crowe, J. E., Jr.; Corti, D.; Virgin, H. W.; Ellebedy, A. H.; Shi, P. Y.; Diamond, M. S., Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat Med* 2021, 27 (4), 717-726.
43. Cherian, S.; Potdar, V.; Jadhav, S.; Yadav, P.; Gupta, N.; Das, M.; Rakshit, P.; Singh, S.; Abraham, P.; Panda, S., Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *bioRxiv* 2021, 2021.04.22.440932.
44. Chien, C. W.; Hou, P. C.; Wu, H. C.; Chang, Y. L.; Lin, S. C.; Lin, S. C.; Lin, B. W.; Lee, J. C.; Chang, Y. J.; Sun, H. S.; Tsai, S. J., Targeting TYRO3 inhibits epithelial-mesenchymal transition and increases drug sensitivity in colon cancer. *Oncogene* 2016, 35 (45), 5872-5881.
45. Childers, M. C.; Daggett, V., Insights from molecular dynamics simulations for computational protein design. *Mol Syst Des Eng* 2017, 2 (1), 9-33.
46. Cock, P. J.; Antao, T.; Chang, J. T.; Chapman, B. A.; Cox, C. J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; de Hoon, M. J., Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 2009, 25 (11), 1422-3.

47. Cohen, P.; Cross, D.; Janne, P. A., Kinase drug discovery 20 years after imatinib: progress and future directions. *Nat Rev Drug Discov* 2021, 20 (7), 551-569.
48. Collier, D. A.; De Marco, A.; Ferreira, I.; Meng, B.; Datir, R. P.; Walls, A. C.; Kemp, S. A.; Bassi, J.; Pinto, D.; Silacci-Fregni, C.; Bianchi, S.; Tortorici, M. A.; Bowen, J.; Culap, K.; Jaconi, S.; Cameroni, E.; Snell, G.; Pizzuto, M. S.; Pellanda, A. F.; Garzoni, C.; Riva, A.; Collaboration, C.-N. B. C.-.; Elmer, A.; Kingston, N.; Graves, B.; McCoy, L. E.; Smith, K. G. C.; Bradley, J. R.; Temperton, N.; Ceron-Gutierrez, L.; Barcenas-Morales, G.; Consortium, C.-G. U.; Harvey, W.; Virgin, H. W.; Lanzavecchia, A.; Piccoli, L.; Doffinger, R.; Wills, M.; Veessler, D.; Corti, D.; Gupta, R. K., Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature* 2021, 593 (7857), 136-141.
49. Collins, M. A.; Ho, J., Accelerating the Calculation of Solute-Solvent Interaction Energies through Systematic Molecular Fragmentation. *J Phys Chem A* 2019, 123 (39), 8476-8484.
50. Colovos, C.; Yeates, T. O., Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 1993, 2 (9), 1511-9.
51. Cowan-Jacob, S. W.; Fendrich, G.; Manley, P. W.; Jahnke, W.; Fabbro, D.; Liebetanz, J.; Meyer, T., The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* 2005, 13 (6), 861-71.
52. Cui, Q.; Sulea, T.; Schrag, J. D.; Munger, C.; Hung, M. N.; Naim, M.; Cygler, M.; Purisima, E. O., Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the MP1-p14 scaffolding complex. *J Mol Biol* 2008, 379 (4), 787-802.
53. Dahiyat, B. I.; Mayo, S. L., De novo protein design: fully automated sequence selection. *Science* 1997, 278 (5335), 82-7.
54. Dai, Y.; Bae, K.; Pampo, C.; Siemann, D. W., Impact of the small molecule Met inhibitor BMS-777607 on the metastatic process in a rodent tumor model with constitutive c-Met activation. *Clin Exp Metastasis* 2012, 29 (3), 253-61.
55. Daina, A.; Michielin, O.; Zoete, V., iLOGP: a simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *J Chem Inf Model* 2014, 54 (12), 3284-301.
56. Daina, A.; Michielin, O.; Zoete, V., SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 2017, 7 (1), 42717.
57. Dallakyan, S.; Olson, A. J., Small-Molecule Library Screening by Docking with PyRx. In *Chemical Biology: Methods and Protocols*, Hempel, J. E.; Williams, C. H.; Hong, C. C., Eds. Springer New York: New York, NY, 2015; pp 243-250.
58. Daniloski, Z.; Jordan, T. X.; Ilmain, J. K.; Guo, X.; Bhabha, G.; tenOever, B. R.; Sanjana, N. E., The Spike D614G mutation increases SARS-CoV-2 infection of multiple human cell types. *Elife* 2021, 10, e65365.

59. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* 1993, 98 (12), 10089-10092.
60. Das Evcimen, N.; King, G. L., The role of protein kinase C activation and the vascular complications of diabetes. *Pharmacol Res* 2007, 55 (6), 498-510.
61. Davies, N. G.; Abbott, S.; Barnard, R. C.; Jarvis, C. I.; Kucharski, A. J.; Munday, J. D.; Pearson, C. A. B.; Russell, T. W.; Tully, D. C.; Washburne, A. D.; Wenseleers, T.; Gimma, A.; Waite, W.; Wong, K. L. M.; van Zandvoort, K.; Silverman, J. D.; Group, C. C.-W.; Consortium, C.-G. U.; Diaz-Ordaz, K.; Keogh, R.; Eggo, R. M.; Funk, S.; Jit, M.; Atkins, K. E.; Edmunds, W. J., Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* 2021, 372 (6538), eabg3055.
62. Davra, V.; Kimani, S. G.; Calianese, D.; Birge, R. B., Ligand Activation of TAM Family Receptors-Implications for Tumor Biology and Therapeutic Response. *Cancers (Basel)* 2016, 8 (12).
63. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A., Role of Molecular Dynamics and Related Methods in Drug Discovery. *J Med Chem* 2016, 59 (9), 4035-61.
64. Dehury, B.; Raina, V.; Misra, N.; Suar, M., Effect of mutation on structure, function and dynamics of receptor binding domain of human SARS-CoV-2 with host cell receptor ACE2: a molecular dynamics simulations study. *J Biomol Struct Dyn* 2021, 39 (18), 7231-7245.
65. Demarest, S. J.; Gardner, J.; Vendel, M. C.; Ailor, E.; Szak, S.; Huang, F.; Doern, A.; Tan, X.; Yang, W.; Grueneberg, D. A.; Richards, E. J.; Endege, W. O.; Harlow, E.; Koopman, L. A., Evaluation of Tyro3 expression, Gas6-mediated Akt phosphorylation, and the impact of anti-Tyro3 antibodies in melanoma cell lines. *Biochemistry* 2013, 52 (18), 3102-18.
66. Deng, T.; Chen, Q.; Han, D., The roles of TAM receptor tyrosine kinases in the mammalian testis and immunoprivileged sites. *Front Biosci (Landmark Ed)* 2016, 21 (2), 316-27.
67. Deng, X.; Garcia-Knight, M. A.; Khalid, M. M.; Servellita, V.; Wang, C.; Morris, M. K.; Sotomayor-Gonzalez, A.; Glasner, D. R.; Reyes, K. R.; Gliwa, A. S.; Reddy, N. P.; Sanchez San Martin, C.; Federman, S.; Cheng, J.; Balcerek, J.; Taylor, J.; Streithorst, J. A.; Miller, S.; Sreekumar, B.; Chen, P. Y.; Schulze-Gahmen, U.; Taha, T. Y.; Hayashi, J. M.; Simoneau, C. R.; Kumar, G. R.; McMahon, S.; Lidsky, P. V.; Xiao, Y.; Hemarajata, P.; Green, N. M.; Espinosa, A.; Kath, C.; Haw, M.; Bell, J.; Hacker, J. K.; Hanson, C.; Wadford, D. A.; Anaya, C.; Ferguson, D.; Frankino, P. A.; Shivram, H.; Lareau, L. F.; Wyman, S. K.; Ott, M.; Andino, R.; Chiu, C. Y., Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell* 2021, 184 (13), 3426-3437 e8.
68. Doncheva, N. T.; Klein, K.; Domingues, F. S.; Albrecht, M., Analyzing and visualizing residue networks of protein structures. *Trends Biochem Sci* 2011, 36 (4), 179-82.
69. Dransfield, I.; Zagorska, A.; Lew, E. D.; Michail, K.; Lemke, G., Mer receptor tyrosine kinase mediates both tethering and phagocytosis of apoptotic cells. *Cell Death Dis* 2015, 6 (2), e1646.

70. Dror, R. O.; Young, C.; Shaw, D. E., Anton, A Special-Purpose Molecular Simulation Machine. In *Encyclopedia of Parallel Computing*, Padua, D., Ed. Springer US: Boston, MA, 2011; pp 60-71.
71. Du, L.; He, Y.; Zhou, Y.; Liu, S.; Zheng, B. J.; Jiang, S., The spike protein of SARS-CoV-a target for vaccine and therapeutic development. *Nat Rev Microbiol* 2009, 7 (3), 226-36.
72. Durrant, J. D.; McCammon, J. A., Molecular dynamics simulations and drug discovery. *BMC Biol* 2011, 9 (1), 71.
73. E, W.; Vanden-Eijnden, E., Transition-path theory and path-finding algorithms for the study of rare events. *Annu Rev Phys Chem* 2010, 61 (1), 391-420.
74. Ekyalongo, R. C.; Mukohara, T.; Funakoshi, Y.; Tomioka, H.; Kataoka, Y.; Shimono, Y.; Chayahara, N.; Toyoda, M.; Kiyota, N.; Minami, H., TYRO3 as a Potential Therapeutic Target in Breast Cancer. *Anticancer Res* 2014, 34 (7), 3337-3345.
75. Ekyalongo, R. C.; Yee, D., Revisiting the IGF-1R as a breast cancer target. *NPJ Precis Oncol* 2017, 1 (1), 1.
76. el Sayadi, H.; Pissaloux, D.; Alberti, L.; Tabone-Eglinger, S.; Ranchere, D.; Decouvlaere, A. V.; Tabone, E.; Ray-Coquard, I.; Caux, C.; Fayette, J.; Blay, J. Y., Autocrine role for Gas6 with Tyro3 and Axl in leiomyosarcomas. *Target Oncol* 2013, 8 (4), 261-9.
77. Endicott, J. A.; Noble, M. E.; Johnson, L. N., The structural basis for control of eukaryotic protein kinases. *Annu Rev Biochem* 2012, 81 (1), 587-613.
78. Ertl, P.; Rohde, B.; Selzer, P., Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem* 2000, 43 (20), 3714-7.
79. Eyal, E.; Bahar, I., Toward a molecular understanding of the anisotropic response of proteins to external forces: insights from elastic network models. *Biophys J* 2008, 94 (9), 3424-35.
80. Fleischmann, W. R., *Viral Genetics*. 4th ed.; University of Texas Medical Branch at Galveston, Galveston (TX): 1996.
81. Fletcher, R.; Powell, M. J. D., A Rapidly Convergent Descent Method for Minimization. *The Computer Journal* 1963, 6 (2), 163-168.
82. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 2004, 47 (7), 1739-49.
83. Gajiwala, K. S.; Grodsky, N.; Bolanos, B.; Feng, J.; Ferre, R.; Timofeevski, S.; Xu, M.; Murray, B. W.; Johnson, T. W.; Stewart, A., The Axl kinase domain in complex with a

- macrocyclic inhibitor offers first structural insights into an active TAM receptor kinase. *J Biol Chem* 2017, 292 (38), 15705-15716.
84. Gan, J. H.; Liu, J. X.; Liu, Y.; Chen, S. W.; Dai, W. T.; Xiao, Z. X.; Cao, Y., DrugRep: an automatic virtual screening server for drug repurposing. *Acta Pharmacol Sin* 2023, 44 (4), 888-896.
 85. Gay, C. M.; Balaji, K.; Byers, L. A., Giving AXL the axe: targeting AXL in human malignancy. *Br J Cancer* 2017, 116 (4), 415-423.
 86. Geng, H.; Chen, F.; Ye, J.; Jiang, F., Applications of Molecular Dynamics Simulation in Structure Prediction of Peptides and Proteins. *Comput Struct Biotechnol J* 2019, 17, 1162-1170.
 87. Geraldles, P.; King, G. L., Activation of protein kinase C isoforms and its impact on diabetic complications. *Circ Res* 2010, 106 (8), 1319-31.
 88. Gilson, M. K.; Honig, B., Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins* 1988, 4 (1), 7-18.
 89. Glassman, C. R.; Tsutsumi, N.; Saxton, R. A.; Lupardus, P. J.; Jude, K. M.; Garcia, K. C., Structure of a Janus kinase cytokine receptor complex reveals the basis for dimeric activation. *Science* 2022, 376 (6589), 163-169.
 90. Gomez, C. E.; Perdiguero, B.; Esteban, M., Emerging SARS-CoV-2 Variants and Impact in Global Vaccination Programs against SARS-CoV-2/COVID-19. *Vaccines (Basel)* 2021, 9 (3).
 91. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., H⁺⁺: a server for estimating pK_as and adding missing hydrogens to macromolecules. *Nucleic Acids Res* 2005, 33 (Web Server issue), W368-71.
 92. Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J Chem Theory Comput* 2012, 8 (5), 1542-1555.
 93. Graham, D. K.; DeRyckere, D.; Davies, K. D.; Earp, H. S., The TAM family: phosphatidylserine sensing receptor tyrosine kinases gone awry in cancer. *Nat Rev Cancer* 2014, 14 (12), 769-85.
 94. Greuber, E. K.; Smith-Pearson, P.; Wang, J.; Pendergast, A. M., Role of ABL family kinases in cancer: from leukaemia to solid tumours. *Nat Rev Cancer* 2013, 13 (8), 559-71.
 95. Grindon, C.; Harris, S.; Evans, T.; Novik, K.; Coveney, P.; Laughton, C., Large-scale molecular dynamics simulation of DNA: implementation and validation of the AMBER98 force field in LAMMPS. *Philos Trans A Math Phys Eng Sci* 2004, 362 (1820), 1373-86.

96. Group, R. C., Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2021, 397 (10285), 1637-1645.
97. Gröllich, C., Cabozantinib: A MET, RET, and VEGFR2 Tyrosine Kinase Inhibitor. In *Small Molecules in Oncology*, Martens, U. M., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 2014; pp 207-214.
98. Gröllich, C., Cabozantinib: Multi-kinase Inhibitor of MET, AXL, RET, and VEGFR2. In *Small Molecules in Oncology*, Martens, U. M., Ed. Springer International Publishing: Cham, 2018; pp 67-75.
99. Gupta, P.; Taiyab, A.; Hassan, M. I., Chapter Three - Emerging role of protein kinases in diabetes mellitus: From mechanism to therapy. In *Advances in Protein Chemistry and Structural Biology*, Donev, R., Ed. Academic Press: 2021; Vol. 124, pp 47-85.
100. Guruprasad, L., Evolutionary relationships and sequence-structure determinants in human SARS coronavirus-2 spike proteins for host receptor recognition. *Proteins* 2020, 88 (11), 1387-1393.
101. Guruprasad, L., Human coronavirus spike protein-host receptor recognition. *Prog Biophys Mol Biol* 2021, 161, 39-53.
102. Guruprasad, L., Human SARS CoV-2 spike protein mutations. *Proteins* 2021, 89 (5), 569-576.
103. Gutte, B., A synthetic 70-amino acid residue analog of ribonuclease S-protein with enzymic activity. *J Biol Chem* 1975, 250 (3), 889-904.
104. Hageman, M. J., Solubility, Solubilization and Dissolution in Drug Delivery During Lead Optimization. In *Optimizing the "Drug-Like" Properties of Leads in Drug Discovery*, Borchardt, R. T.; Kerns, E. H.; Hageman, M. J.; Thakker, D. R.; Stevens, J. L., Eds. Springer New York: New York, NY, 2006; pp 99-130.
105. Harding, H. P.; Calfon, M.; Urano, F.; Novoa, I.; Ron, D., Transcriptional and translational control in the Mammalian unfolded protein response. *Annu Rev Cell Dev Biol* 2002, 18 (1), 575-99.
106. Harrigan, M. P.; Sultan, M. M.; Hernandez, C. X.; Husic, B. E.; Eastman, P.; Schwantes, C. R.; Beauchamp, K. A.; McGibbon, R. T.; Pande, V. S., MSMBuilder: Statistical Models for Biomolecular Dynamics. *Biophys J* 2017, 112 (1), 10-15.
107. Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; Del Rio, J. F.; Wiebe, M.; Peterson, P.; Gerard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E., Array programming with NumPy. *Nature* 2020, 585 (7825), 357-362.
108. Harvey, W. T.; Carabelli, A. M.; Jackson, B.; Gupta, R. K.; Thomson, E. C.; Harrison, E. M.; Ludden, C.; Reeve, R.; Rambaut, A.; Consortium, C.-G. U.; Peacock, S. J.; Robertson,

- D. L., SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* 2021, 19 (7), 409-424.
109. Heinzlmeir, S.; Kudlinzki, D.; Sreeramulu, S.; Klaeger, S.; Gande, S. L.; Linhard, V.; Wilhelm, M.; Qiao, H.; Helm, D.; Ruprecht, B.; Saxena, K.; Medard, G.; Schwalbe, H.; Kuster, B., Chemical Proteomics and Structural Biology Define EPHA2 Inhibition by Clinical Kinase Drugs. *ACS Chem Biol* 2016, 11 (12), 3400-3411.
 110. Herum, K. M.; Lunde, I. G.; McCulloch, A. D.; Christensen, G., The Soft- and Hard-Heartedness of Cardiac Fibroblasts: Mechanotransduction Signaling Pathways in Fibrosis of the Heart. *J Clin Med* 2017, 6 (5).
 111. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* 1997, 18 (12), 1463-1472.
 112. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* 2008, 4 (3), 435-47.
 113. Hinton, G. E.; Salakhutdinov, R. R., Reducing the dimensionality of data with neural networks. *Science* 2006, 313 (5786), 504-7.
 114. Hollingshead, M. G.; Greenberg, N.; Gottholm-Ahalt, M.; Camalier, R.; Johnson, B. C.; Collins, J. M.; Doroshow, J. H., ROADMAPS: An Online Database of Response Data, Dosing Regimens, and Toxicities of Approved Oncology Drugs as Single Agents to Guide Preclinical In Vivo Studies. *Cancer Res* 2022, 82 (12), 2219-2225.
 115. Homeyer, N.; Gohlke, H., Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Mol Inform* 2012, 31 (2), 114-22.
 116. Horby, P. W.; Pessoa-Amorim, G.; Peto, L.; Brightling, C. E.; Sarkar, R.; Thomas, K.; Jeebun, V.; Ashish, A.; Tully, R.; Chadwick, D.; Sharafat, M.; Stewart, R.; Rudran, B.; Baillie, J. K.; Buch, M. H.; Chappell, L. C.; Day, J. N.; Furst, S. N.; Jaki, T.; Jeffery, K.; Juszczak, E.; Lim, W. S.; Montgomery, A.; Mumford, A.; Rowan, K.; Thwaites, G.; Mafham, M.; Haynes, R.; Landray, M. J., Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): preliminary results of a randomised, controlled, open-label, platform trial. *medRxiv* 2021, 2021.02.11.21249258.
 117. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 2006, 65 (3), 712-25.
 118. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* 2011, 51 (1), 69-82.

119. Hu, J.; Ahuja, L. G.; Meharena, H. S.; Kannan, N.; Kornev, A. P.; Taylor, S. S.; Shaw, A. S., Kinase regulation by hydrophobic spine assembly in cancer. *Mol Cell Biol* 2015, 35 (1), 264-76.
120. Huang, Y.; Yang, C.; Xu, X. F.; Xu, W.; Liu, S. W., Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. *Acta Pharmacol Sin* 2020, 41 (9), 1141-1149.
121. Huey, M. G.; Minson, K. A.; Earp, H. S.; DeRyckere, D.; Graham, D. K., Targeting the TAM Receptors in Leukemia. *Cancers (Basel)* 2016, 8 (11), 1-22.
122. Hunter, J. D., Matplotlib: A 2D graphics environment. *Comput Sci Eng* 2007, 9 (3), 90-95.
123. Hunter, T.; Cooper, J. A., Protein-tyrosine kinases. *Annu Rev Biochem* 1985, 54 (1), 897-930.
124. Husic, B. E.; Pande, V. S., Markov State Models: From an Art to a Science. *J Am Chem Soc* 2018, 140 (7), 2386-2396.
125. Irwin, J. J.; Shoichet, B. K.; Mysinger, M. M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y., Automated docking screens: a feasibility study. *J Med Chem* 2009, 52 (18), 5712-20.
126. Istifli, E. S.; Netz, P. A.; Sihoglu Tepe, A.; Sarikurkcu, C.; Tepe, B., Understanding the molecular interaction of SARS-CoV-2 spike mutants with ACE2 (angiotensin converting enzyme 2). *J Biomol Struct Dyn* 2022, 40 (23), 12760-12771.
127. Jia, X., Solvation Free Energy Calculations: The Combination between the Implicitly Polarized Fixed-charge Model and the Reference Potential Strategy. *J Comput Chem* 2019, 40 (32), 2801-2809.
128. Jiang, T.; Liu, Z.; Liu, W.; Chen, J.; Zheng, Z.; Duan, M., The Conformational Transition Pathways and Hidden Intermediates in DFG-Flip Process of c-Met Kinase Revealed by Metadynamics Simulations. *J Chem Inf Model* 2022, 62 (15), 3651-3663.
129. Jo, S.; Kim, T.; Iyer, V. G.; Im, W., CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 2008, 29 (11), 1859-65.
130. John, B.; Sali, A., Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res* 2003, 31 (14), 3982-92.
131. Johnson, L. N.; Noble, M. E.; Owen, D. J., Active and inactive protein kinases: structural basis for regulation. *Cell* 1996, 85 (2), 149-58.
132. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 1983, 79 (2), 926-935.
133. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 1996, 118 (45), 11225-11236.

134. Juyoux, P.; Galdadas, I.; Gobbo, D.; Velsen, J. v.; Pelosse, M.; Tully, M.; Vadas, O.; Gervasio, F. L.; Pellegrini, E.; Bowler, M. W., Architecture of the MKK6-p38 α complex defines the basis of MAPK specificity and activation. *bioRxiv* 2023, 2022.07.04.498667.
135. Kasaian, K.; Wiseman, S. M.; Walker, B. A.; Schein, J. E.; Hirst, M.; Moore, R. A.; Mungall, A. J.; Marra, M. A.; Jones, S. J., Putative BRAF activating fusion in a medullary thyroid cancer. *Cold Spring Harb Mol Case Stud* 2016, 2 (2), a000729.
136. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J., The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015, 10 (6), 845-58.
137. Khan, A.; Zia, T.; Suleman, M.; Khan, T.; Ali, S. S.; Abbasi, A. A.; Mohammad, A.; Wei, D. Q., Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *J Cell Physiol* 2021, 236 (10), 7045-7057.
138. Kim, J.; Ahuja, L. G.; Chao, F. A.; Xia, Y.; McClendon, C. L.; Kornev, A. P.; Taylor, S. S.; Veglia, G., A dynamic hydrophobic core orchestrates allostery in protein kinases. *Sci Adv* 2017, 3 (4), e1600663.
139. Kimani, S. G.; Kumar, S.; Davra, V.; Chang, Y. J.; Kasikara, C.; Geng, K.; Tsou, W. I.; Wang, S.; Hoque, M.; Bohac, A.; Lewis-Antes, A.; De Lorenzo, M. S.; Kotenko, S. V.; Birge, R. B., Normalization of TAM post-receptor signaling reveals a cell invasive signature for Axl tyrosine kinase. *Cell Commun Signal* 2016, 14 (1), 19.
140. Knubel, K. H.; Pernu, B. M.; Sufit, A.; Nelson, S.; Pierce, A. M.; Keating, A. K., MerTK inhibition is a novel therapeutic approach for glioblastoma multiforme. *Oncotarget* 2014, 5 (5), 1338-51.
141. Koes, D. R.; Camacho, C. J., ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic Acids Res* 2012, 40 (Web Server issue), W409-14.
142. 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., 3rd, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 2000, 33 (12), 889-97.
143. Kopp, J.; Schwede, T., The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res* 2004, 32 (Database issue), D230-4.
144. Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; Hastie, K. M.; Parker, M. D.; Partridge, D. G.; Evans, C. M.; Freeman, T. M.; de Silva, T. I.; Sheffield, C.-G. G.; McDanal, C.; Perez, L. G.; Tang, H.; Moon-Walker, A.; Whelan, S. P.; LaBranche, C. C.; Saphire, E. O.; Montefiori, D. C., Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* 2020, 182 (4), 812-827 e19.

145. Kornev, A. P.; Haste, N. M.; Taylor, S. S.; Eyck, L. F., Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A* 2006, 103 (47), 17783-8.
146. Kramer, B.; Rarey, M.; Lengauer, T., Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins* 1999, 37 (2), 228-41.
147. Kufareva, I.; Abagyan, R., Methods of protein structure comparison. *Methods Mol Biol* 2012, 857, 231-57.
148. Kunitomo, R.; Bajorath, J., Combining Similarity Searching and Network Analysis for the Identification of Active Compounds. *Acs Omega* 2018, 3 (4), 3768-3777.
149. Kyriakidis, N. C.; Lopez-Cortes, A.; Gonzalez, E. V.; Grimaldos, A. B.; Prado, E. O., SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates. *NPJ Vaccines* 2021, 6 (1), 28.
150. Lacy, S.; Nielsen, J.; Yang, B.; Miles, D.; Nguyen, L.; Hutmacher, M., Population exposure-response analysis of cabozantinib efficacy and safety endpoints in patients with renal cell carcinoma. *Cancer Chemother Pharmacol* 2018, 81 (6), 1061-1070.
151. Lai, S.; Safaei, J.; Pelech, S., Evolutionary Ancestry of Eukaryotic Protein Kinases and Choline Kinases. *J Biol Chem* 2016, 291 (10), 5199-205.
152. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X., Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 2020, 581 (7807), 215-220.
153. Lane, T. J.; Bowman, G. R.; Beauchamp, K.; Voelz, V. A.; Pande, V. S., Markov state model reveals folding and functional dynamics in ultra-long MD trajectories. *J Am Chem Soc* 2011, 133 (45), 18413-9.
154. Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D., DOCK 6: combining techniques to model RNA-small molecule complexes. *Rna* 2009, 15 (6), 1219-30.
155. Laskowski, R. A.; Jablonska, J.; Pravda, L.; Varekova, R. S.; Thornton, J. M., PDBsum: Structural summaries of PDB entries. *Protein Sci* 2018, 27 (1), 129-134.
156. Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M., Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Crystallogr* 1993, 26 (2), 283-291.
157. Lee, P. Y.; Yeoh, Y.; Low, T. Y., A recent update on small-molecule kinase inhibitors for targeted cancer therapy and their therapeutic insights from mass spectrometry-based proteomic analysis. *FEBS J* 2023, 290 (11), 2845-2864.
158. Lemke, G., Biology of the TAM receptors. *Cold Spring Harb Perspect Biol* 2013, 5 (11), a009076.

159. Leone, G. M.; Candido, S.; Lavoro, A.; Vivarelli, S.; Gattuso, G.; Calina, D.; Libra, M.; Falzone, L. Clinical Relevance of Targeted Therapy and Immune-Checkpoint Inhibition in Lung Cancer Pharmaceuticals 2023, 15(4), p.1252
160. Levinson, N. M.; Kuchment, O.; Shen, K.; Young, M. A.; Koldobskiy, M.; Karplus, M.; Cole, P. A.; Kuriyan, J., A Src-like inactive conformation in the abl tyrosine kinase domain. PLoS Biol 2006, 4 (5), e144.
161. Levitt, M.; Lifson, S., Refinement of protein conformations using a macromolecular energy minimization procedure. J Mol Biol 1969, 46 (2), 269-79.
162. Lew, E. D.; Oh, J.; Burrola, P. G.; Lax, I.; Zagorska, A.; Traves, P. G.; Schlessinger, J.; Lemke, G., Differential TAM receptor-ligand-phospholipid interactions delimit differential TAM bioactivities. Elife 2014, 3.
163. Li, R.; Liu, J.; Zhang, H., The challenge of emerging SARS-CoV-2 mutants to vaccine development. J Genet Genomics 2021, 48 (2), 102-106.
164. Li, W.; Moore, M. J.; Vasilieva, N.; Sui, J.; Wong, S. K.; Berne, M. A.; Somasundaran, M.; Sullivan, J. L.; Luzuriaga, K.; Greenough, T. C.; Choe, H.; Farzan, M., Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 2003, 426 (6965), 450-4.
165. Li, Y.; Dong, Z., Effect of Clustering Algorithm on Establishing Markov State Model for Molecular Dynamics Simulations. J Chem Inf Model 2016, 56 (6), 1205-15.
166. Likas, A.; Vlassis, N.; Verbeek, J. J., The global k-means clustering algorithm. Pattern Recogn 2003, 36 (2), 451-461.
167. Lill, M., Virtual Screening in Drug Design. In In Silico Models for Drug Discovery, Kortagere, S., Ed. Humana Press: Totowa, NJ, 2013; pp 1-12.
168. Lill, M. A.; Thompson, J. J., Solvent interaction energy calculations on molecular dynamics trajectories: increasing the efficiency using systematic frame selection. J Chem Inf Model 2011, 51 (10), 2680-9.
169. Lin, F.-Y.; MacKerell, A. D., Force Fields for Small Molecules. In Biomolecular Simulations: Methods and Protocols, Bonomi, M.; Camilloni, C., Eds. Springer New York: New York, NY, 2019; pp 21-54.
170. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E., Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 2010, 78 (8), 1950-8.
171. Lipinski, C. A., Lead- and drug-like compounds: the rule-of-five revolution. Drug Discov Today Technol 2004, 1 (4), 337-41.

172. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001, 46 (1-3), 3-26.
173. Luan, B.; Wang, H.; Huynh, T., Enhanced binding of the N501Y-mutated SARS-CoV-2 spike protein to the human ACE2 receptor: insights from molecular dynamics simulations. *FEBS Lett* 2021, 595 (10), 1454-1461.
174. Lubin, J. H.; Zardecki, C.; Dolan, E. M.; Lu, C.; Shen, Z.; Dutta, S.; Westbrook, J. D.; Hudson, B. P.; Goodsell, D. S.; Williams, J. K.; Voigt, M.; Sarma, V.; Xie, L.; Venkatachalam, T.; Arnold, S.; Alfaro Alvarado, L. H.; Catalfano, K.; Khan, A.; McCarthy, E.; Staggers, S.; Tinsley, B.; Trudeau, A.; Singh, J.; Whitmore, L.; Zheng, H.; Benedek, M.; Currier, J.; Dresel, M.; Duvvuru, A.; Dyszel, B.; Fingar, E.; Hennen, E. M.; Kirsch, M.; Khan, A. A.; Labrie-Cleary, C.; Laporte, S.; Lenkeit, E.; Martin, K.; Orellana, M.; Ortiz-Alvarez de la Campa, M.; Paredes, I.; Wheeler, B.; Rupert, A.; Sam, A.; See, K.; Soto Zapata, S.; Craig, P. A.; Hall, B. L.; Jiang, J.; Koeppe, J. R.; Mills, S. A.; Pikaart, M. J.; Roberts, R.; Bromberg, Y.; Hoyer, J. S.; Duffy, S.; Tischfield, J.; Ruiz, F. X.; Arnold, E.; Baum, J.; Sandberg, J.; Brannigan, G.; Khare, S. D.; Burley, S. K., Evolution of the SARS-CoV-2 proteome in three dimensions (3D) during the first 6 months of the COVID-19 pandemic. *Proteins* 2022, 90 (5), 1054-1080.
175. Lui, A.; Vanleuven, J.; Perekopskiy, D.; Liu, D.; Xu, D.; Alzayat, O.; Elgokhy, T.; Do, T.; Gann, M.; Martin, R.; Liu, D.-Z. FDA-Approved Kinase Inhibitors in Preclinical and Clinical Trials for Neurological Disorders *Pharmaceuticals* 2022, 15(12), p.1546
176. Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R., Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res* 2022, 50 (W1), W276-W279.
177. Madhusudhan, M.; Marti-Renom, M. A.; Eswar, N.; John, B.; Pieper, U.; Karchin, R.; Shen, M.-Y.; Sali, A., Comparative protein structure modeling. In *The proteomics protocols handbook*, Springer: 2005; pp 831-860.
178. Mahgoub, R. E.; Atatreh, N.; Ghattas, M. A., Chapter Three - Using filters in virtual screening: A comprehensive guide to minimize errors and maximize efficiency. In *Annual Reports in Medicinal Chemistry*, Caballero, J., Ed. Academic Press: 2022; Vol. 59, pp 99-136.
179. Maia, E. H. B.; Assis, L. C.; de Oliveira, T. A.; da Silva, A. M.; Taranto, A. G., Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Front Chem* 2020, 8, 343.
180. Malolepsza, E.; Strodel, B.; Khalili, M.; Trygubenko, S.; Fejer, S. N.; Wales, D. J., Symmetrization of the AMBER and CHARMM force fields. *J Comput Chem* 2010, 31 (7), 1402-9.
181. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S., The protein kinase complement of the human genome. *Science* 2002, 298 (5600), 1912-34.

182. Mark, P.; Nilsson, L., Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A* 2001, 105 (43), 9954-9960.
183. Markosian, C.; Staquicini, D. I.; Dogra, P.; Dodero-Rojas, E.; Lubin, J. H.; Tang, F. H. F.; Smith, T. L.; Contessoto, V. G.; Libutti, S. K.; Wang, Z.; Cristini, V.; Khare, S. D.; Whitford, P. C.; Burley, S. K.; Onuchic, J. N.; Pasqualini, R.; Arap, W., Genetic and Structural Analysis of SARS-CoV-2 Spike Protein for Universal Epitope Selection. *Mol Biol Evol* 2022, 39 (5), msac091.
184. Markowitz, J. N.; Fancher, K. M., Cabozantinib: A Multitargeted Oral Tyrosine Kinase Inhibitor. *Pharmacotherapy* 2018, 38 (3), 357-369.
185. McCammon, J. A.; Gelin, B. R.; Karplus, M., Dynamics of folded proteins. *Nature* 1977, 267 (5612), 585-90.
186. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernandez, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S., MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys J* 2015, 109 (8), 1528-32.
187. McKinney, W. In *Data structures for statistical computing in python*, Proceedings of the 9th Python in Science Conference, Austin, TX: 2010; pp 51-56.
188. Meagher, K. L.; Redman, L. T.; Carlson, H. A., Development of polyphosphate parameters for use with the AMBER force field. *J Comput Chem* 2003, 24 (9), 1016-25.
189. Meireles, L.; Gur, M.; Bakan, A.; Bahar, I., Pre-existing soft modes of motion uniquely defined by native contact topology facilitate ligand binding to proteins. *Protein Sci* 2011, 20 (10), 1645-58.
190. Meng, Y.; Roux, B., Locking the active conformation of c-Src kinase through the phosphorylation of the activation loop. *J Mol Biol* 2014, 426 (2), 423-35.
191. Metzner, P.; Schutte, C.; Vanden-Eijnden, E., Transition Path Theory for Markov Jump Processes. *Multiscale Model Sim* 2009, 7 (3), 1192-1219.
192. Meyer, A. S.; Zweemer, A. J.; Lauffenburger, D. A., The AXL Receptor is a Sensor of Ligand Spatial Heterogeneity. *Cell Syst* 2015, 1 (1), 25-36.
193. Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* 2012, 8 (9), 3314-21.
194. Miller, W. T., Tyrosine kinase signaling and the emergence of multicellularity. *Biochim Biophys Acta* 2012, 1823 (6), 1053-7.
195. Mohammadi, E.; Shafiee, F.; Shahzamani, K.; Ranjbar, M. M.; Alibakhshi, A.; Ahangarzadeh, S.; Beikmohammadi, L.; Shariati, L.; Hooshmandi, S.; Ataei, B.;

- Javanmard, S. H., Novel and emerging mutations of SARS-CoV-2: Biomedical implications. *Biomed Pharmacother* 2021, 139, 111599.
196. Mohanty, S.; Oruganty, K.; Kwon, A.; Byrne, D. P.; Ferries, S.; Ruan, Z.; Hanold, L. E.; Katiyar, S.; Kennedy, E. J.; Eysers, P. A.; Kannan, N., Hydrophobic Core Variations Provide a Structural Framework for Tyrosine Kinase Evolution and Functional Specialization. *PLoS Genet* 2016, 12 (2), e1005885.
 197. Mokaya, M.; Imrie, F.; van Hoorn, W. P.; Kalisz, A.; Bradley, A. R.; Deane, C. M., Testing the limits of SMILES-based de novo molecular generation with curriculum and deep reinforcement learning. *Nat Mach Intell* 2023, 5 (4), 386.
 198. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 2009, 30 (16), 2785-91.
 199. Morris, J. H.; Huang, C. C.; Babbitt, P. C.; Ferrin, T. E., structureViz: linking Cytoscape and UCSF Chimera. *Bioinformatics* 2007, 23 (17), 2345-7.
 200. Muhammed, M. T.; Aki-Yalcin, E., Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des* 2019, 93 (1), 12-20.
 201. Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A., A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* 2006, 63 (4), 928-38.
 202. Myers, K. V.; Amend, S. R.; Pienta, K. J., Targeting Tyro3, Axl and MerTK (TAM receptors): implications for macrophages in the tumor microenvironment. *Mol Cancer* 2019, 18 (1), 94.
 203. Nagamalla, L.; Kumar, J. V. S.; Shaik, M. R.; Sanjay, C.; Alsamhan, A. M.; Kasim, M. A.; Alwarthan, A. Identification of Novel AXL Kinase Inhibitors Using Ligand-Based Pharmacophore Screening and Molecular Dynamics Simulations Crystals [Online], 2022.
 204. Naim, M.; Bhat, S.; Rankin, K. N.; Dennis, S.; Chowdhury, S. F.; Siddiqi, I.; Drabik, P.; Sulea, T.; Bayly, C. I.; Jakalian, A.; Purisima, E. O., Solvated interaction energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space. *J Chem Inf Model* 2007, 47 (1), 122-33.
 205. Nandipati, K. C.; Subramanian, S.; Agrawal, D. K., Protein kinases: mechanisms and downstream targets in inflammation-mediated obesity and insulin resistance. *Mol Cell Biochem* 2017, 426 (1-2), 27-45.
 206. Nelson, M. T.; Humphrey, W.; Gursoy, A.; Dalke, A.; Kale, L. V.; Skeel, R. D.; Schulten, K., NAMD: A parallel, object oriented molecular dynamics program. *Int J Supercomput Ap* 1996, 10 (4), 251-268.
 207. Nguyen, K. Q.; Tsou, W. I.; Calarese, D. A.; Kimani, S. G.; Singh, S.; Hsieh, S.; Liu, Y.; Lu, B.; Wu, Y.; Garforth, S. J.; Almo, S. C.; Kotenko, S. V.; Birge, R. B., Overexpression

- of MERTK receptor tyrosine kinase in epithelial cancer cells drives efferocytosis in a gain-of-function capacity. *J Biol Chem* 2014, 289 (37), 25737-49.
208. Noe, F.; Clementi, C., Kinetic distance and kinetic maps from molecular dynamics simulation. *J Chem Theory Comput* 2015, 11 (10), 5002-11.
 209. Noh, J. Y.; Jeong, H. W.; Shin, E. C., SARS-CoV-2 mutations, vaccines, and immunity: implication of variants of concern. *Signal Transduct Target Ther* 2021, 6 (1), 203.
 210. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An open chemical toolbox. *J Cheminform* 2011, 3 (1), 33.
 211. Ozono, S.; Zhang, Y.; Ode, H.; Sano, K.; Tan, T. S.; Imai, K.; Miyoshi, K.; Kishigami, S.; Ueno, T.; Iwatani, Y.; Suzuki, T.; Tokunaga, K., SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. *Nat Commun* 2021, 12 (1), 848.
 212. Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K., Forces contributing to the conformational stability of proteins. *FASEB J* 1996, 10 (1), 75-83.
 213. Pantano, F.; Ribelli, G.; Iuliani, M.; Fioramonti, M.; Lealos, M.; Zoccoli, A.; Santini, D.; Muto, G.; Tonini, G., New Molecular Targets in Metastatic Prostate Cancer. *Journal of Cancer Therapy* 2016, 07 (06), 388-401.
 214. Paolino, M.; Choidas, A.; Wallner, S.; Pranjic, B.; Uribesalga, I.; Loeser, S.; Jamieson, A. M.; Langdon, W. Y.; Ikeda, F.; Fededa, J. P.; Cronin, S. J.; Nitsch, R.; Schultz-Fademrecht, C.; Eickhoff, J.; Menninger, S.; Unger, A.; Torka, R.; Gruber, T.; Hinterleitner, R.; Baier, G.; Wolf, D.; Ullrich, A.; Klebl, B. M.; Penninger, J. M., The E3 ligase Cbl-b and TAM receptors regulate cancer metastasis via natural killer cells. *Nature* 2014, 507 (7493), 508-12.
 215. Parenti, M. D.; Rastelli, G., Advances and applications of binding affinity prediction methods in drug discovery. *Biotechnol Adv* 2012, 30 (1), 244-50.
 216. Parrinello, M.; Rahman, A., Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. *J Appl Phys* 1981, 52 (12), 7182-7190.
 217. Parsons, S. J.; Parsons, J. T., Src family kinases, key regulators of signal transduction. *Oncogene* 2004, 23 (48), 7906-9.
 218. Patone, M.; Thomas, K.; Hatch, R.; Tan, P. S.; Coupland, C.; Liao, W.; Mouncey, P.; Harrison, D.; Rowan, K.; Horby, P.; Watkinson, P.; Hippisley-Cox, J., Mortality and critical care unit admission associated with the SARS-CoV-2 lineage B.1.1.7 in England: an observational cohort study. *Lancet Infect Dis* 2021, 21 (11), 1518-1528.
 219. Patterson, H.; Nibbs, R.; McInnes, I.; Siebert, S., Protein kinase inhibitors in the treatment of inflammatory and autoimmune diseases. *Clin Exp Immunol* 2014, 176 (1), 1-10.

220. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E., Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 2011, 12, 2825-2830.
221. Perez, E.; Lindblad, J. L.; Bergmann, A., Tumor-promoting function of apoptotic caspases by an amplification loop involving ROS, macrophages and JNK in *Drosophila*. *Elife* 2017, 6, e26747.
222. Perez, F.; Granger, B. E., IPython: A system for interactive scientific computing. *Comput Sci Eng* 2007, 9 (3), 21-29.
223. Perez-Hernandez, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noe, F., Identification of slow molecular order parameters for Markov model construction. *J Chem Phys* 2013, 139 (1), 015102.
224. Phillips, D. C., The three-dimensional structure of an enzyme molecule. *Sci Am* 1966, 215 (5), 78-90.
225. Pinato, D. J.; Chowdhury, S.; Stebbing, J., TAMing resistance to multi-targeted kinase inhibitors through Axl and Met inhibition. *Oncogene* 2016, 35 (21), 2684-6.
226. Pinzi, L.; Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery *International Journal of Molecular Sciences*, 2019, 20(18), p.4331.
227. Piovesan, D.; Minervini, G.; Tosatto, S. C., The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res* 2016, 44 (W1), W367-74.
228. Plowman, G. D.; Sudarsanam, S.; Bingham, J.; Whyte, D.; Hunter, T., The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999, 96 (24), 13603-10.
229. Potts, R. O.; Guy, R. H., Predicting skin permeability. *Pharm Res* 1992, 9 (5), 663-9.
230. Prinz, J. H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schutte, C.; Noe, F., Markov models of molecular kinetics: generation and validation. *J Chem Phys* 2011, 134 (17), 174105.
231. Qin, S.; Li, A.; Yi, M.; Yu, S.; Zhang, M.; Wu, K., Recent advances on anti-angiogenesis receptor tyrosine kinase inhibitors in cancer therapy. *J Hematol Oncol* 2019, 12 (1), 27.
232. Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V., Stereochemistry of polypeptide chain configurations. *J Mol Biol* 1963, 7 (1), 95-9.
233. Ramachandran, G. N.; Venkatachalam, C. M., Stereochemical criteria for polypeptides and proteins. IV. Standard dimensions for the cis-peptide unit and conformation of cis-polypeptides. *Biopolymers* 1968, 6 (9), 1255-62.

234. Rambaut, A.; Holmes, E. C.; O'Toole, A.; Hill, V.; McCrone, J. T.; Ruis, C.; du Plessis, L.; Pybus, O. G., A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020, 5 (11), 1403-1407.
235. Rankin, E. B.; Giaccia, A. J., The Receptor Tyrosine Kinase AXL in Cancer Progression. *Cancers (Basel)* 2016, 8 (11).
236. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M., Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations. *J Am Chem Soc* 1992, 114 (25), 10024-10035.
237. Reback, J.; McKinney, W.; Van Den Bossche, J.; Augspurger, T.; Cloud, P.; Klein, A.; Hawkins, S.; Roeschke, M.; Tratner, J.; She, C., pandas-dev/pandas: Pandas 1.0. 5. Zenodo 2020.
238. Rees-Spear, C.; Muir, L.; Griffith, S. A.; Heaney, J.; Aldon, Y.; Snitselaar, J. L.; Thomas, P.; Graham, C.; Seow, J.; Lee, N.; Rosa, A.; Roustan, C.; Houlihan, C. F.; Sanders, R. W.; Gupta, R. K.; Cherepanov, P.; Stauss, H. J.; Nastouli, E.; Investigators, S.; Doores, K. J.; van Gils, M. J.; McCoy, L. E., The effect of spike mutations on SARS-CoV-2 neutralization. *Cell Rep* 2021, 34 (12), 108890.
239. Robert, X.; Gouet, P., Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 2014, 42 (Web Server issue), W320-4.
240. Roberts, R. J.; Vincze, T.; Posfai, J.; Macelis, D., REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 2015, 43 (Database issue), D298-9.
241. Robinson, R., Confirming the importance of the R-spine: new insights into protein kinase regulation. *PLoS Biol* 2013, 11 (10), e1001681.
242. Roskoski, R., Jr., Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. *Pharmacol Res* 2015, 94, 9-25.
243. Roskoski, R., Jr., Properties of FDA-approved small molecule protein kinase inhibitors: A 2022 update. *Pharmacol Res* 2022, 175, 106037.
244. Roskoski, R., Jr., Rule of five violations among the FDA-approved small molecule protein kinase inhibitors. *Pharmacol Res* 2023, 191, 106774.
245. Roskoski, R., Jr., Properties of FDA-approved small molecule protein kinase inhibitors: A 2023 update. *Pharmacol Res* 2023, 187, 106552.
246. Rothlin, C. V.; Carrera-Silva, E. A.; Bosurgi, L.; Ghosh, S., TAM receptor signaling in immune homeostasis. *Annu Rev Immunol* 2015, 33, 355-91.
247. Ruder, S., An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747* 2016.

248. Ruiz-Morales, J. M.; Heng, D. Y., Cabozantinib in the treatment of advanced renal cell carcinoma: clinical trial evidence and experience. *Ther Adv Urol* 2016, 8 (6), 338-347.
249. Sali, A., Comparative protein modeling by satisfaction of spatial restraints. *Mol Med Today* 1995, 1 (6), 270-7.
250. Sali, A.; Blundell, T. L., Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993, 234 (3), 779-815.
251. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C., An overview of the Amber biomolecular simulation package. *Wires Comput Mol Sci* 2013, 3 (2), 198-210.
252. Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput* 2013, 9 (9), 3878-88.
253. Salvalaglio, M.; Tiwary, P.; Parrinello, M., Assessing the Reliability of the Dynamics Reconstructed from Metadynamics. *J Chem Theory Comput* 2014, 10 (4), 1420-5.
254. Sasaki, T.; Knyazev, P. G.; Clout, N. J.; Cheburkin, Y.; Gohring, W.; Ullrich, A.; Timpl, R.; Hohenester, E., Structural basis for Gas6-Axl signalling. *EMBO J* 2006, 25 (1), 80-7.
255. Sausville, E. A.; Johnson, J. I., Molecules for the millennium: how will they look? New drug discovery year 2000. *Br J Cancer* 2000, 83 (11), 1401-4.
256. Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Perez-Hernandez, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J. H.; Noe, F., PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J Chem Theory Comput* 2015, 11 (11), 5525-42.
257. Schmitz, R.; Valls, A. F.; Yerbes, R.; von Richter, S.; Kahlert, C.; Loges, S.; Weitz, J.; Schneider, M.; Ruiz de Almodovar, C.; Ulrich, A.; Schmidt, T., TAM receptors Tyro3 and Mer as novel targets in colorectal cancer. *Oncotarget* 2016, 7 (35), 56355-56370.
258. Schwantes, C. R.; Pande, V. S., Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *J Chem Theory Comput* 2013, 9 (4), 2000-2009.
259. Segal, M. R., Machine learning benchmarks and random forest regression. 2004.
260. Seif, F.; Khoshmirsafa, M.; Aazami, H.; Mohsenzadegan, M.; Sedighi, G.; Bahar, M., The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Commun Signal* 2017, 15 (1), 23.
261. Shah, M.; Ahmad, B.; Choi, S.; Woo, H. G., Mutations in the SARS-CoV-2 spike RBD are responsible for stronger ACE2 binding and poor anti-SARS-CoV mAbs cross-neutralization. *Comput Struct Biotechnol J* 2020, 18, 3402-3414.
262. Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F., Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* 2020, 117 (21), 11727-11734.

263. Sharma, J.; Kumar Bhardwaj, V.; Singh, R.; Rajendran, V.; Purohit, R.; Kumar, S., An in-silico evaluation of different bioactive molecules of tea for their inhibition potency against non structural protein-15 of SARS-CoV-2. *Food Chem* 2021, 346, 128933.
264. Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, C. R.; Ierardi, D. J.; Kolossvary, I.; Klepeis, J. L.; Layman, T.; McLeavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y. B.; Spengler, J.; Theobald, M.; Towles, B.; Wang, S. C., Anton, a Special-Purpose Machine for Molecular Dynamics Simulation. *Conf Proc Int Symp C* 2007, 35 (2), 1-12.
265. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W., Atomic-level characterization of the structural dynamics of proteins. *Science* 2010, 330 (6002), 341-6.
266. Shen, M. Y.; Sali, A., Statistical potential for assessment and prediction of protein structures. *Protein Sci* 2006, 15 (11), 2507-24.
267. Shi, J.; Blundell, T. L.; Mizuguchi, K., FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol* 2001, 310 (1), 243-57.
268. Shukla, D.; Meng, Y.; Roux, B.; Pande, V. S., Activation pathway of Src kinase reveals intermediate states as targets for drug design. *Nat Commun* 2014, 5 (1), 3397.
269. Sievers, F.; Higgins, D. G., Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* 2018, 27 (1), 135-145.
270. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011, 7 (1), 539.
271. Singh, I.; Singh, S.; Verma, V.; Uversky, V. N.; Chandra, R., In silico evaluation of the resistance of the T790M variant of epidermal growth factor receptor kinase to cancer drug Erlotinib. *J Biomol Struct Dyn* 2018, 36 (16), 4209-4219.
272. Singh, N.; Chaput, L.; Villoutreix, B. O., Virtual screening web servers: designing chemical probes and drug candidates in the cyberspace. *Brief Bioinform* 2021, 22 (2), 1790-1818.
273. Sitkoff, D.; Lockhart, D. J.; Sharp, K. A.; Honig, B., Calculation of electrostatic effects at the amino terminus of an alpha helix. *Biophysical journal* 1994, 67 (6), 2251-2260.
274. Sitkoff, D.; Sharp, K. A.; Honig, B., Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. *J Phys Chem-Us* 1994, 98 (7), 1978-1988.
275. Skora, L.; Mestan, J.; Fabbro, D.; Jahnke, W.; Grzesiek, S., 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* 2013, 110 (47), E4437-45.

276. Smith, W.; Yong, C. W.; Rodger, P. M., DL_POLY: Application to molecular simulation. *Mol Simulat* 2002, 28 (5), 385-471.
277. Sousa da Silva, A. W.; Vranken, W. F., ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res Notes* 2012, 5, 367.
278. Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A., Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate–DNA Helices. *J Am Chem Soc* 1998, 120 (37), 9401-9409.
279. Straatsma, T. P.; McCammon, J. A., Free energy evaluation from molecular dynamics simulations using force fields including electronics polarization. *Chemical Physics Letters* 1991, 177 (4-5), 433-440.
280. Sulea, T.; Purisima, E. O., The Solvated Interaction Energy Method for Scoring Binding Affinities. In *Computational Drug Discovery and Design*, Baron, R., Ed. Springer New York: New York, NY, 2012; pp 295-303.
281. Sultan, M. M.; Denny, R. A.; Unwalla, R.; Lovering, F.; Pande, V. S., Millisecond dynamics of BTK reveal kinome-wide conformational plasticity within the apo kinase domain. *Sci Rep* 2017, 7 (1), 15604.
282. Sultan, M. M.; Kiss, G.; Pande, V. S., Towards simple kinetic models of functional dynamics for a kinase subfamily. *Nat Chem* 2018, 10 (9), 903-909.
283. Sun, H.; Jin, Z.; Yang, C.; Akkermans, R. L.; Robertson, S. H.; Spenley, N. A.; Miller, S.; Todd, S. M., COMPASS II: extended coverage for polymer and drug-like molecule databases. *J Mol Model* 2016, 22 (2), 47.
284. Sun, H.; Ren, P.; Fried, J. R., The COMPASS force field: parameterization and validation for phosphazenes. *Comput Theor Polym S* 1998, 8 (1-2), 229-246.
285. Sun, Y.; Sun, L.; An, Y.; Shen, X., Cabozantinib, a Novel c-Met Inhibitor, Inhibits Colorectal Cancer Development in a Xenograft Model. *Med Sci Monit* 2015, 21, 2316-21.
286. Sunseri, J.; Koes, D. R., Pharmit: interactive exploration of chemical space. *Nucleic Acids Res* 2016, 44 (W1), W442-8.
287. Taka, E.; Yilmaz, S. Z.; Golcuk, M.; Kilinc, C.; Aktas, U.; Yildiz, A.; Gur, M., Critical Interactions Between the SARS-CoV-2 Spike Glycoprotein and the Human ACE2 Receptor. *J Phys Chem B* 2021, 125 (21), 5537-5548.
288. Taylor, S. S.; Kornev, A. P., Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci* 2011, 36 (2), 65-77.
289. Tian, F.; Tong, B.; Sun, L.; Shi, S.; Zheng, B.; Wang, Z.; Dong, X.; Zheng, P., N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. *Elife* 2021, 10, e69091.

290. Tian, Z.; Niu, X.; Yao, W., Receptor Tyrosine Kinases in Osteosarcoma Treatment: Which Is the Key Target? *Frontiers in Oncology* 2020, 10.
291. Tong, J. C.; Ranganathan, S., *Computer-aided vaccine design*. Elsevier: 2013.
292. Torres, P. H. M.; Sodero, A. C. R.; Jofily, P.; Silva-Jr, F. P. Key Topics in Molecular Docking for Drug Design *International Journal of Molecular Sciences* [Online], 2019. https://mdpi-res.com/d_attachment/ijms/ijms-20-04574/article_deploy/ijms-20-04574-v2.pdf?version=1568616656.
293. Tridente, G., Chapter 21 - Cabozantinib. In *Adverse Events and Oncotargeted Kinase Inhibitors*, Tridente, G., Ed. Academic Press: 2017; pp 457-480.
294. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010, 31 (2), 455-61.
295. Tsou, W. I.; Nguyen, K. Q.; Calarese, D. A.; Garforth, S. J.; Antes, A. L.; Smirnov, S. V.; Almo, S. C.; Birge, R. B.; Kotenko, S. V., Receptor tyrosine kinases, TYRO3, AXL, and MER, demonstrate distinct patterns and complex regulation of ligand-induced activation. *J Biol Chem* 2014, 289 (37), 25750-63.
296. Turner, N. A.; Blythe, N. M., Cardiac Fibroblast p38 MAPK: A Critical Regulator of Myocardial Remodeling. *J Cardiovasc Dev Dis* 2019, 6 (3).
297. Uras, F.; Kucuk, B.; Bingol Ozakpinar, O.; Demir, A. M., Growth Arrest-Specific 6 (Gas6) and TAM Receptors in Mouse Platelets. *Turk J Haematol* 2015, 32 (1), 58-63.
298. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J., GROMACS: fast, flexible, and free. *J Comput Chem* 2005, 26 (16), 1701-18.
299. van Gunsteren, W. F.; Berendsen, H. J., Thermodynamic cycle integration by computer simulation as a tool for obtaining free energy differences in molecular chemistry. *J Comput Aided Mol Des* 1987, 1 (2), 171-6.
300. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Zidek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S., AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* 2022, 50 (D1), D439-D444.
301. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., Improved protein-ligand docking using GOLD. *Proteins* 2003, 52 (4), 609-23.
302. Verlet, L., Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* 1967, 159 (1), 98-103.

303. Verlet, L., Computer "Experiments" on Classical Fluids. II. Equilibrium Correlation Functions. *Physical Review* 1968, 165 (1), 201-214.
304. Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, I.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy, C., SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020, 17 (3), 261-272.
305. Vouri, M.; Hafizi, S., TAM Receptor Tyrosine Kinases in Cancer Drug Resistance. *Cancer Res* 2017, 77 (11), 2775-2778.
306. Wan, S.; Yan, R.; Jiang, Y.; Li, Z.; Zhang, J.; Wu, X., Insight into binding mechanisms of EGFR allosteric inhibitors using molecular dynamics simulations and free energy calculations. *J Biomol Struct Dyn* 2019, 37 (16), 4384-4394.
307. Wang, C.; Greene, D.; Xiao, L.; Qi, R.; Luo, R., Recent Developments and Applications of the MMPBSA Method. *Front Mol Biosci* 2017, 4 (87), 87.
308. Wang, C.; Jin, H.; Wang, N.; Fan, S.; Wang, Y.; Zhang, Y.; Wei, L.; Tao, X.; Gu, D.; Zhao, F.; Fang, J.; Yao, M.; Qin, W., Gas6/Axl Axis Contributes to Chemoresistance and Metastasis in Breast Cancer through Akt/GSK-3 β / β -catenin Signaling. *Theranostics* 2016, 6 (8), 1205-19.
309. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 2006, 25 (2), 247-60.
310. Wang, L.; Zheng, G.; Liu, X.; Ni, D.; He, X.; Cheng, J.; Lu, S., Molecular dynamics simulations provide insights into the origin of gleevec's selectivity toward human tyrosine kinases. *J Biomol Struct Dyn* 2019, 37 (10), 2733-2744.
311. Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J., Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* 2020, 181 (4), 894-904 e9.
312. Wang, W. B.; Liang, Y.; Jin, Y. Q.; Zhang, J.; Su, J. G.; Li, Q. M., E484K mutation in SARS-CoV-2 RBD enhances binding affinity with hACE2 but reduces interactions with neutralizing antibodies and nanobodies: Binding free energy calculation studies. *J Mol Graph Model* 2021, 109, 108035.
313. Watanabe, T.; Imamura, T.; Hiasa, Y., Roles of protein kinase R in cancer: Potential as a therapeutic target. *Cancer Sci* 2018, 109 (4), 919-925.
314. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL:

- homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018, 46 (W1), W296-W303.
315. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* 2014, 47 (1), 5 6 1-32.
316. Weiner, P. K.; Kollman, P. A., Amber - Assisted Model-Building with Energy Refinement - a General Program for Modeling Molecules and Their Interactions. *J Comput Chem* 1981, 2 (3), 287-303.
317. Whitman, S. P.; Kohlschmidt, J.; Maharry, K.; Volinia, S.; Mrozek, K.; Nicolet, D.; Schwind, S.; Becker, H.; Metzeler, K. H.; Mendler, J. H.; Eisfeld, A. K.; Carroll, A. J.; Powell, B. L.; Carter, T. H.; Baer, M. R.; Kolitz, J. E.; Park, I. K.; Stone, R. M.; Caligiuri, M. A.; Marcucci, G.; Bloomfield, C. D., GAS6 expression identifies high-risk adult AML patients: potential implications for therapy. *Leukemia* 2014, 28 (6), 1252-1258.
318. Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V., Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J Chem Theory Comput* 2014, 10 (3), 1228-1241.
319. Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z., A new coronavirus associated with human respiratory disease in China. *Nature* 2020, 579 (7798), 265-269.
320. Wu, G.; Ma, Z.; Cheng, Y.; Hu, W.; Deng, C.; Jiang, S.; Li, T.; Chen, F.; Yang, Y., Targeting Gas6/TAM in cancer cells and tumor microenvironment. *Mol Cancer* 2018, 17 (1), 20.
321. Wu, G.; Ma, Z.; Hu, W.; Wang, D.; Gong, B.; Fan, C.; Jiang, S.; Li, T.; Gao, J.; Yang, Y., Molecular insights of Gas6/TAM in cancer development and therapy. *Cell Death Dis* 2017, 8 (3), e2700.
322. Wu, G.; Robertson, D. H.; Brooks, C. L., 3rd; Vieth, M., Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm. *J Comput Chem* 2003, 24 (13), 1549-62.
323. Wu, H.; Mardt, A.; Pasquali, L.; Noe, F., Deep Generative Markov State Models. *Adv Neur In* 2018, 31, 3979–3988.
324. Wu, Q.; Qian, W.; Sun, X.; Jiang, S., Small-molecule inhibitors, immune checkpoint inhibitors, and more: FDA-approved novel therapeutic drugs for solid tumors from 1991 to 2021. *J Hematol Oncol* 2022, 15 (1), 143.
325. Xiao, T.; Lu, J.; Zhang, J.; Johnson, R. I.; McKay, L. G. A.; Storm, N.; Lavine, C. L.; Peng, H.; Cai, Y.; Rits-Volloch, S.; Lu, S.; Quinlan, B. D.; Farzan, M.; Seaman, M. S.; Griffiths, A.; Chen, B., A trimeric human angiotensin-converting enzyme 2 as an anti-SARS-CoV-2 agent. *Nat Struct Mol Biol* 2021, 28 (2), 202-209.

326. Xie, Z.; Lee, Y. H.; Boeke, M.; Jilaveanu, L. B.; Liu, Z.; Bottaro, D. P.; Kluger, H. M.; Shuch, B., MET Inhibition in Clear Cell Renal Cell Carcinoma. *J Cancer* 2016, 7 (10), 1205-14.
327. Xu, C.; Wang, Y.; Liu, C.; Zhang, C.; Han, W.; Hong, X.; Wang, Y.; Hong, Q.; Wang, S.; Zhao, Q.; Wang, Y.; Yang, Y.; Chen, K.; Zheng, W.; Kong, L.; Wang, F.; Zuo, Q.; Huang, Z.; Cong, Y., Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Sci Adv* 2021, 7 (1), eabe5575.
328. Yang, L. W.; Eyal, E.; Bahar, I.; Kitao, A., Principal component analysis of native ensembles of biomolecular structures (PCA_NEST): insights into functional dynamics. *Bioinformatics* 2009, 25 (5), 606-14.
329. Yang, T. J.; Yu, P. Y.; Chang, Y. C.; Liang, K. H.; Tso, H. C.; Ho, M. R.; Chen, W. Y.; Lin, H. T.; Wu, H. C.; Hsu, S. D., Effect of SARS-CoV-2 B.1.1.7 mutations on spike protein structure and function. *Nat Struct Mol Biol* 2021, 28 (9), 731-739.
330. Yang, Z.; Lasker, K.; Schneidman-Duhovny, D.; Webb, B.; Huang, C. C.; Pettersen, E. F.; Goddard, T. D.; Meng, E. C.; Sali, A.; Ferrin, T. E., UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J Struct Biol* 2012, 179 (3), 269-78.
331. Zagorska, A.; Traves, P. G.; Lew, E. D.; Dransfield, I.; Lemke, G., Diversification of TAM receptor tyrosine kinase function. *Nat Immunol* 2014, 15 (10), 920-8.
332. Zhang, L.; Jackson, C. B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B. D.; Rangarajan, E. S.; Pan, A.; Vanderheiden, A.; Suthar, M. S.; Li, W.; Izard, T.; Rader, C.; Farzan, M.; Choe, H., SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat Commun* 2020, 11 (1), 6013.
333. Zhang, Q.; Wu, X.; Zhang, H.; Wu, Q.; Fu, M.; Hua, L.; Zhu, X.; Guo, Y.; Zhang, L.; You, Q.; Wang, L., Protein Phosphatase 5-Recruiting Chimeras for Accelerating Apoptosis-Signal-Regulated Kinase 1 Dephosphorylation with Antiproliferative Activity. *J Am Chem Soc* 2023, 145 (2), 1118-1128.
334. Zhang, Y.; He, X.; Zhai, J.; Ji, B.; Man, V. H.; Wang, J., In silico binding profile characterization of SARS-CoV-2 spike protein and its mutants bound to human ACE2 receptor. *Brief Bioinform* 2021, 22 (6), bbab188.
335. Zhang, Y.; Li, N.; Chen, Q.; Yan, K.; Liu, Z.; Zhang, X.; Liu, P.; Chen, Y.; Han, D., Breakdown of immune homeostasis in the testis of mice lacking Tyro3, Axl and Mer receptor tyrosine kinases. *Immunol Cell Biol* 2013, 91 (6), 416-26.
336. Zhao, H.; Caflisch, A., Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics. *Bioorg Med Chem Lett* 2013, 23 (20), 5721-6.
337. Zhou, D.; Dejnirattisai, W.; Supasa, P.; Liu, C.; Mentzer, A. J.; Ginn, H. M.; Zhao, Y.; Duyvesteyn, H. M. E.; Tuekprakhon, A.; Nutalai, R.; Wang, B.; Paesen, G. C.; Lopez-

- Camacho, C.; Slon-Campos, J.; Hallis, B.; Coombes, N.; Bewley, K.; Charlton, S.; Walter, T. S.; Skelly, D.; Lumley, S. F.; Dold, C.; Levin, R.; Dong, T.; Pollard, A. J.; Knight, J. C.; Crook, D.; Lambe, T.; Clutterbuck, E.; Bibi, S.; Flaxman, A.; Bittaye, M.; Belij-Rammerstorfer, S.; Gilbert, S.; James, W.; Carroll, M. W.; Klenerman, P.; Barnes, E.; Dunachie, S. J.; Fry, E. E.; Mongkolsapaya, J.; Ren, J.; Stuart, D. I.; Screaton, G. R., Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* 2021, 184 (9), 2348-2361 e6.
338. Zhou, L.; Liu, X. D.; Sun, M.; Zhang, X.; German, P.; Bai, S.; Ding, Z.; Tannir, N.; Wood, C. G.; Matin, S. F.; Karam, J. A.; Tamboli, P.; Sircar, K.; Rao, P.; Rankin, E. B.; Laird, D. A.; Hoang, A. G.; Walker, C. L.; Giaccia, A. J.; Jonasch, E., Targeting MET and AXL overcomes resistance to sunitinib therapy in renal cell carcinoma. *Oncogene* 2016, 35 (21), 2687-97.
 339. Zhu, C.; Wei, Y.; Wei, X., AXL receptor tyrosine kinase as a promising anti-cancer approach: functions, molecular mechanisms and clinical applications. *Mol Cancer* 2019, 18 (1), 153.
 340. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W.; China Novel Coronavirus, I.; Research, T., A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020, 382 (8), 727-733.
 341. Zwanzig, R. W., High-temperature equation of state by a perturbation method. I. Nonpolar gases. *The Journal of Chemical Physics* 1954, 22 (8), 1420-1426.
 342. Zwanzig, R. W., High-temperature equation of state by a perturbation method. II. polar gases. *The Journal of Chemical Physics* 1955, 23 (10), 1915-1922.

Published research articles

a. Enhanced metastable state models of TAM kinase binding to cabozantinib explains the dynamic nature of receptor tyrosine kinases

Gatta K. R. S. Naresh & Lalitha Guruprasad* (2020)

Journal of Biomolecular Structure and Dynamics, DOI:10.1080/07391102.2020.1730968

b. Mutations in the receptor-binding domain of human SARS Cov-2 spike protein increasing its affinity to bind human ACE-2 receptor

Gatta K. R. S. Naresh & Lalitha Guruprasad* (2022)

Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2022.2032354

c. Dynamic conformational states of apo, ATP and cabozantinib bound TAM kinases to differentiate active-inactive kinetic models

Gatta K. R. S. Naresh & Lalitha Guruprasad* (2023)

Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2022.2162128

e. In silico design, modelling and molecular mechanisms of Axl receptor tyrosine kinase inhibitors (manuscript under preparation)

Book chapter:

f. Chapter 8 - Structure-based methods in drug design. In Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development.

Guruprasad, L*.; Andola, P.; Banerjee, A.; Laxman, D.; Naresh, G. K. R. S.,(2023)

Academic Press. DOI:10.1016/B978-0-443-18638-7.00003-7

Conferences participation

- “**Data Science in Biomedical Research using R**” held on 19th - 23th Dec 2018 (CMSD) at University of Hyderabad.
- “**Numerical Methods and Computer Programming**” lecture series conducted by “UGC-DAE Consortium for Scientific Research” (NMCP2018) held on 6th - 10th Aug 2018 Indore.
- “**Machine Learning for Science (ML4Science)**” held 29-30 Nov 2019 at IIIT Hyderabad.
- “**Science and Technology Entrepreneurship Program (STEP)**” Jointly organized by ERC, OIA, MSU and Research and Innovation, Technology and Entrepreneurship (RITE) cell University of Hyderabad held from 7th to 17th Jul 2021.
- “**International Conference On frontier Areas of Science and Technology**” conducted by “12th India- Japan Science and Technology Conclave” (ICFAST- 2022) held on 09-10th Sep 2022 at University of Hyderabad Campus, Hyderabad India.
- “**Statistical Machine Learning for Biologists**” held on 19th - 22th Dec 2022 (CMSD) at University of Hyderabad.
- I participated one day workshop of “**ACS ON CAMPUS**” conducted by American Chemical Society team held on 13th Apr 2023 (CMSD) at University of Hyderabad.
- I participated three day workshop of “**Hands on Workshop on Microbiome and Genome Analysis**” conducted by DBT -Centre for Microbial Informatics team held on 19th -21st Jul 2023 (CMSD) at University of Hyderabad.

Poster presentations

- “**Structural & Conformational Analysis of TAM RTKs with Cabozantinib**” at “15th Annual In-House Symposium (ChemFest-2018)” held on Mar 9-10, 2018 (national conference) at University of Hyderabad.
- “**TAM Kinase Structure and Dynamics for Inhibitor Design**” at “16th Annual In-House Symposium (ChemFest-2019)” held on 22-23 Feb 2019 (national conference) at University of Hyderabad.

- **“Structural Dynamics of Cabozantinib bound TAM receptor tyrosine kinases”** conducted by “International Conference on Drug Discovery 2020 (ICDD-2020) held on 29th Feb – 2nd Mar 2020 at BITS Pilani Campus Hyderabad.
- **“Dynamic States of ATP and inhibitor bound receptor tyrosine kinase to explore open and closed states”** at “17th Annual In-House Symposium (ChemFest-2022)” held on 22-23 April 2022 (national conference) at University of Hyderabad.
- **“Dynamic active and inactive states of TAM RTK kinase domain binding with ATP and inhibitor cabozantinib”** conducted by “International Conference on Drug Discovery 2020 (ICDD 2022) held on 10th Nov – 11th Nov 2022 at BITS Pilani KK Birla Goa Campus, Goa.

Oral presentations

- **“Structural Dynamics of ATP and Inhibitor bound receptor tyrosine kinases to explore open and closed states”** conducted by “Emerging Infectious disease & Therapeutics strategies (MedChem 2021) held on 1st - 3rd Dec 2021 at IIT Madras Chennai.
- **“Dynamic States of ATP and inhibitor bound receptor tyrosine kinase to explore open and closed states”** at “16th Annual In-House Symposium (ChemFest-2022)” held on 22-23 April 2022 (national conference) at University of Hyderabad.
- **“Enhanced metastable dynamic states of TAM kinases for inhibitor binding”** prePhD Seminar at School of Chemistry University of Hyderabad (May 2023).



Log in

| Register



Cart

Journal of Biomolecular Structure and Dynamics >

Volume 39, 2021 - Issue 4

332 | 3

Views | CrossRef citations to date | 1 | Altmetric

Research Articles

Enhanced metastable state models of TAM kinase binding to cabozantinib explains the dynamic nature of receptor tyrosine kinases

Gatta K. R. S. Naresh & Lalitha Guruprasad

Pages 1213-1235 | Received 02 Nov 2019, Accepted 03 Feb 2020, Accepted author version posted online: 19 Feb 2020, Published online: 09 Mar 2020

Download citation <https://doi.org/10.1080/07391102.2020.1730968>



Full Article

Figures & data

References

Supplemental

Citations

Metrics

Reprints & Permissions

Read this article

Abstract

Receptor tyrosine kinases (RTKs) are essential proteins in the regulation of cell signaling. Tyro3, Axl and Mer are members of TAM RTKs and are overexpressed in several cancer forms. Kinase inhibitors such as cabozantinib, foretinib are reported to inhibit TAM kinases at nanomolar concentrations. The atomistic details of structure and mechanism of functional regulation is required to understand their normal physiological process and when bound to an inhibitor. The docking of cabozantinib into the active state conformations of TAM kinases (crystal structure and computational models) revealed the best binding pose and the complex formation that is mediated through non-bonding interactions involving the hinge region residues. The alterations in the conformations and the regions of flexibility in apo and complexed TAM kinases as a course of time are studied using 250 ns molecular dynamics (MD) simulations. The post-MD trajectory analysis using Python libraries like ProDy, MDTrai

101 Views | 0 CrossRef citations to date | 1 Altmetric

Research Article

Dynamic conformational states of apo, ATP and cabozantinib bound TAM kinases to differentiate active-inactive kinetic models

Gatta K. R. S. Naresh  & Lalitha Guruprasad  

Received 14 Sep 2022, Accepted 18 Dec 2022, Published online: 02 Jan 2023

 Download citation
  <https://doi.org/10.1080/07391102.2022.2162128>



[Full Article](#) |
 [Figures & data](#) |
 [References](#) |
 [Supplemental](#) |
 [Citations](#) |
 [Metrics](#)

[Reprints & Permissions](#) |
 [Read this article](#)

Abstract

The dynamically active and inactive conformations of kinases play a crucial role in the activation of intracellular downstream signaling pathways. The all-atom molecular dynamics (MD) simulations at microsecond (μ s) timescale and longer provide robust insights into the structural details of conformational alterations in kinases that contribute to their cellular metabolic activities and signaling pathways. Tyro3, Axl and Mer (TAM) receptor tyrosine kinases (RTKs) are overexpressed in several types of human cancers. Cabozantinib, a small molecule inhibitor constrains the activity of TAM kinases at nanomolar concentrations. The apo, complexes of ATP (active state) and cabozantinib (active and inactive states) with TAM RTKs were studied by 1 μ s MD simulations followed by trajectory analyses. The dynamic mechanistic pathways intrinsic to the kinase activity and protein conformational landscape in the cabozantinib bound TAM kinases are revealed due to the alterations in the P-loop. α -helix

173 | 2 | 0
 Views | CrossRef citations to date | Altmetric

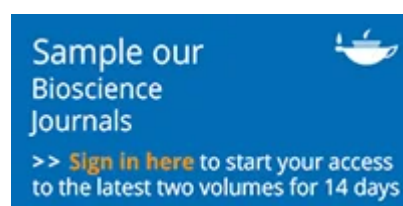
Research Articles

Mutations in the receptor-binding domain of human SARS CoV-2 spike protein increases its affinity to bind human ACE-2 receptor

Gatta K. R. S. Naresh & Lalitha Guruprasad 

Pages 2368-2381 | Received 22 Oct 2021, Accepted 12 Jan 2022, Published online: 03 Feb 2022

 Download citation
  <https://doi.org/10.1080/07391102.2022.2032354>



- [Full Article](#) |
 [Figures & data](#) |
 [References](#) |
 [Supplemental](#) |
 [Citations](#) |
 [Metrics](#)
- [Reprints & Permissions](#) |
 Read this article

Abstract

The severe acute respiratory syndrome virus-2 (SARS CoV-2) infection has resulted in the current global pandemic. The binding of SARS CoV-2 spike protein receptor-binding domain (RBD) to the human angiotensin converting enzyme-2 (ACE-2) receptor causes the host infection. The spike protein has undergone several mutations with reference to the initial strain isolated during December 2019 from Wuhan, China. A number of these mutant strains have been reported as variants of concern and as variants being monitored. Some of these mutants are known to be responsible for increased transmissibility of the virus. The reason for the increased transmissibility caused by the point mutations can be understood by studying the structural implications and inter-molecular interactions in the binding of viral spike protein RBD and human ACE-2. Here, we use the crystal structure of the RBD in complex with ACE-2 available in the public domain and analyse the 250 ns molecular



Chapter 8 - Structure-based methods in drug design

[Lalitha Guruprasad](#), [Priyanka Andola](#), [Adrija Banerjee](#), [Durgam Laxman](#), [Gatta K.R.S. Naresh](#)

School of Chemistry, University of Hyderabad, Hyderabad, India

Available online 5 June 2023, Version of Record 5 June 2023.

Show less ^

Outline | Share | Cite

<https://doi.org/10.1016/B978-0-443-18638-7.00003-7>

[Get rights and content](#)

Abstract

Drug discovery using traditional methods is expensive on resources. Advances in algorithms for computational methods toward solving problems in chemical biology and affordable computational costs have made a rapid revolution in the last three decades resulting in the discovery of novel drugs for targeted therapy using rationale approaches. Structure-based drug design (SBDD) is a popular methodology employed when the structure of the target receptor is known. For a chemical molecule to receive United States Food and Drug Administration (FDA) approval to be marketed as a safe drug, it should have desired biological effects, pass all the clinical trials for its assessment of efficacy and safety, and the chemistry should be novel enough to be patentable. SBDD methods initially design hit molecules that are based on receptor-ligand complementarity. Improvisation of the hit to lead molecules is a cyclical process that comprises the computational design of chemical molecules, synthetic chemistry, experimental validation, and further modifications till an optimized drug candidate that shows the therapeutic effect is discovered.

< Previous

Next >

Keywords

Structure-based drug design; Protein structure; Binding site identification; Virtual screening; Molecular docking; Fragment-based drug design; De novo drug design; Covalent inhibitors; Multiparameter optimization

[Recommended articles](#)

Cited by (0)

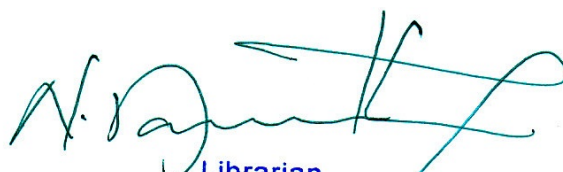
[View full text](#)

Copyright © 2023 Elsevier Inc. All rights reserved.



Enhanced metastable dynamic states of TAM kinases for inhibitor binding

by Gatta Krs Naresh



Librarian
Indira Gandhi Memorial Library
UNIVERSITY OF HYDERABAD
Central University P.O.
HYDERABAD-500 046.

Submission date: 28-Jul-2023 05:37PM (UTC+0530)

Submission ID: 2138021000

File name: Gatta_KRS_Naresh.pdf (12.08M)

Word count: 40630

Character count: 223605

ORIGINALITY REPORT

48%

SIMILARITY INDEX

23%

INTERNET SOURCES

48%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1

Gatta K. R. S. Naresh, Lalitha Guruprasad.
"Enhanced metastable state models of TAM
kinase binding to cabozantinib explains the
dynamic nature of receptor tyrosine kinases",
Journal of Biomolecular Structure and
Dynamics, 2020

Publication

16%

Lalitha Guruprasad

Professor
School of Chemistry
University of Hyderabad
Hyderabad-500 046.

2

www.biorxiv.org

Internet Source

Student Paper

Lalitha Guruprasad

Professor
School of Chemistry
University of Hyderabad
Hyderabad-500 046.

14%

3

Gatta K. R. S. Naresh, Lalitha Guruprasad.
"Mutations in the receptor-binding domain of
human SARS CoV-2 spike protein increases its
affinity to bind human ACE-2 receptor",
Journal of Biomolecular Structure and
Dynamics, 2022

Publication

14%

Lalitha Guruprasad

Professor
School of Chemistry
University of Hyderabad
Hyderabad-500 046.

4

www.researchgate.net

Internet Source

1%

5

Lalitha Guruprasad, K R S NARESH GATTA.
"Dynamic conformational states of apo and

1%

Lalitha Guruprasad

Professor
School of Chemistry
University of Hyderabad
Hyderabad-500 046.

cabozantinib bound TAM kinases to
differentiate active-inactive kinetic models",
Cold Spring Harbor Laboratory, 2021

Publication

6

Submitted to University of Hyderabad,
Hyderabad

Student Paper

<1 %

7

www.ukessays.com

Internet Source

<1 %

8

Priyanka Andola, Jishu Pagag, Durgam
Laxman, Lalitha Guruprasad. "Fragment-
based inhibitor design for SARS-CoV2 main
protease", Structural Chemistry, 2022

Publication

<1 %

9

"Drug Repurposing for Emerging Infectious
Diseases and Cancer", Springer Science and
Business Media LLC, 2023

Publication

<1 %

10

Laxman Durgam, Lalitha Guruprasad.
"Molecular mechanism of ATP and RNA
binding to Zika virus NS3 helicase and
identification of repurposed drugs using
molecular dynamics simulations", Journal of
Biomolecular Structure and Dynamics, 2021

Publication

<1 %

11

core.ac.uk

Internet Source

<1 %

12	repository.dl.itc.u-tokyo.ac.jp Internet Source	<1 %
13	"Structural Biology in Drug Discovery", Wiley, 2020 Publication	<1 %
14	"Handbook of Materials Modeling", Springer Science and Business Media LLC, 2020 Publication	<1 %
15	www.creative-diagnostics.com Internet Source	<1 %
16	Kaifu Gao, Rui Wang, Jiahui Chen, Limei Cheng et al. "Methodology-Centered Review of Molecular Modeling, Simulation, and Prediction of SARS-CoV-2", Chemical Reviews, 2022 Publication	<1 %
17	sites.google.com Internet Source	<1 %
18	www.nature.com Internet Source	<1 %
19	Methods in Molecular Biology, 2012. Publication	<1 %
20	wit.com.tw Internet Source	<1 %

21 Maaged Abdullah, Lalitha Guruprasad. <1 %
"Computational fragment-based design of
Wee1 kinase inhibitors with tricyclic core
scaffolds", Structural Chemistry, 2018
Publication

22 ebin.pub <1 %
Internet Source

23 Shalini Mathpal, Priyanka Sharma, Tushar
Joshi, Veena Pande et al. "Identification of
Zinc-Binding Inhibitors of Matrix
Metalloproteinase-9 to Prevent Cancer
Through Deep Learning and Molecular
Dynamics Simulation Approach", Frontiers in
Molecular Biosciences, 2022
Publication

24 Wan, Xiaobo, Yue Ma, Christopher L.
McClendon, Lily Jun-shen Huang, and Niu
Huang. "Ab Initio Modeling and Experimental
Assessment of Janus Kinase 2 (JAK2) Kinase-
Pseudokinase Complex Structure", PLoS
Computational Biology, 2013.
Publication

25 mdpi-res.com <1 %
Internet Source

26 www.drugs.com <1 %
Internet Source

27

Eileen Socher, Lukas Heger, Friedrich Paulsen, Friederike Zunke, Philipp Arnold. "Molecular dynamics simulations of the delta and omicron SARS-CoV-2 spike – ACE2 complexes reveal distinct changes between both variants", Computational and Structural Biotechnology Journal, 2022

Publication

<1 %

28

Submitted to University of Witwatersrand

Student Paper

<1 %

29

Wei Ge, Limin Wang, Ji Xu, Feiguo Chen, Guangzheng Zhou, Liqiang Lu, Qi Chang, Jinghai Li. "Discrete simulation of granular and particle-fluid flows: from fundamental study to engineering application", Reviews in Chemical Engineering, 2017

Publication

<1 %

Exclude quotes On

Exclude matches

< 14 words

Exclude bibliography On