Chapter 4: Identification of human tau-tubulin kinase 1 inhibitors

4.1   Introduction

Protein kinases play vital role in cellular activation processes, especially in regulation of cell signaling pathways and thus serve as valuable therapeutic target for numerous diseases [1]. Over 30 kinase inhibitors have gained the market for different physiological disorders *viz.* cancers, fibrosis, and arthritis [2]. The lack of kinase inhibitors, for the treatment of neurodegenerative disorders, offers an interesting opportunity for drug development. In normal brain, tau phosphoprotein contains 2–3 moles phosphate per mole of protein. In AD brain, it is abnormally hyperphosphorylated by 2–3 folds at various sites *i.e.,* Thr181, Ser195, Ser198, Ser199, Ser202, Thr205, Thr212, Ser214, Thr217, Thr231, Ser235, Ser262, Ser353, Ser396, Ser400, Ser404, Ser409, and Ser422 [3, 4].

Tau-tubulin kinase (TTBK), a serine/threonine and tyrosine kinase enzyme of casein kinase 1(CK-1) superfamily, is implicated in various physiological processes including mitosis, ciliogenesis, and neurotransmission, *etc*. The enzyme has two isoforms: TTBK1 and TTBK2, consisting of highly homologous catalytic domains with different non-catalytic domains. TTBK1, a neuron specific tau kinase, has molecular weight of 142.7 kDa, but the full-length protein migrates up to 230 kDa, possibly due to the negatively charged polyglutamate region. It is responsible for phosphorylation (both Mg^{2+}- and Mn^{2+}-dependent manner) of tau protein and paired helical filaments (PHFs) [5, 6]. The phosphorylations take place at residues Ser198, Ser199, Ser202, and Ser422, to facilitate tau aggregation and increase the formation of NFTs in AD [7].

TTBK1 is expressed only in central nervous system (CNS) and induces Cdk5/GSK3β (cyclin-dependent kinase 5/glycogen synthase kinase 3β) activation [8], tau aggregation [6, 7], and axonal degeneration [9]. TTBK2 is widely distributed, not only in brain but also in heart, muscle, liver, thymus, spleen, lung, kidney, testis, and ovaries [10]. It is also attributed to tau phosphorylation on residues Ser208 and Ser210 [11], which are characteristics of PHFs phosphorylation in brain. Phosphorylation of TTBK1 and 2 is also linked to neurodegenerative diseases like AD, ALS, and SCA type 11 [11]. Hence, design and development of selective TTBK1 inhibitors are expected to be fruitful and possibly lead to safe drug(s) for the treatment of neurodegenerative diseases.

The structure of TTBK1 resembles other protein kinases with an enriched β-strand N-terminal domain and an α-helical C-terminal domain. An extended 'hinge' region (residues 108–111) connects both the terminals. The P-loop (residues 40–49) is part of N-terminal domain, while Asp-Phe-Gly (DFG) motif (residues 176–178) and flexible activation loop (residues 178–202) are fragments of C-terminal (**Figure 4.1A**) [11]. The superimposed TTBK1-cocrystal ligand, 2KC, complex (PDB4NFN), and apoprotein structure (PDB4NFM), appear equivalent in conformation, as well as configuration (**Figure 4.1B**).

Three reported inhibitors of TTBK1 have strengthened TTBK1 as a new target for the treatment of neurodegenerative diseases and their inhibitors as potential drugs. Energy-optimized structure-based pharmacophore (e-pharmacophore) approach, followed by virtual screening of free 'ZINC15' database molecules with molecular dynamics (MD) simulation assisted us in identifying new TTBK1 inhibitors. The workflow of structure-based hit(s) identification is straightforward and depicted in **Figure 4.2**.
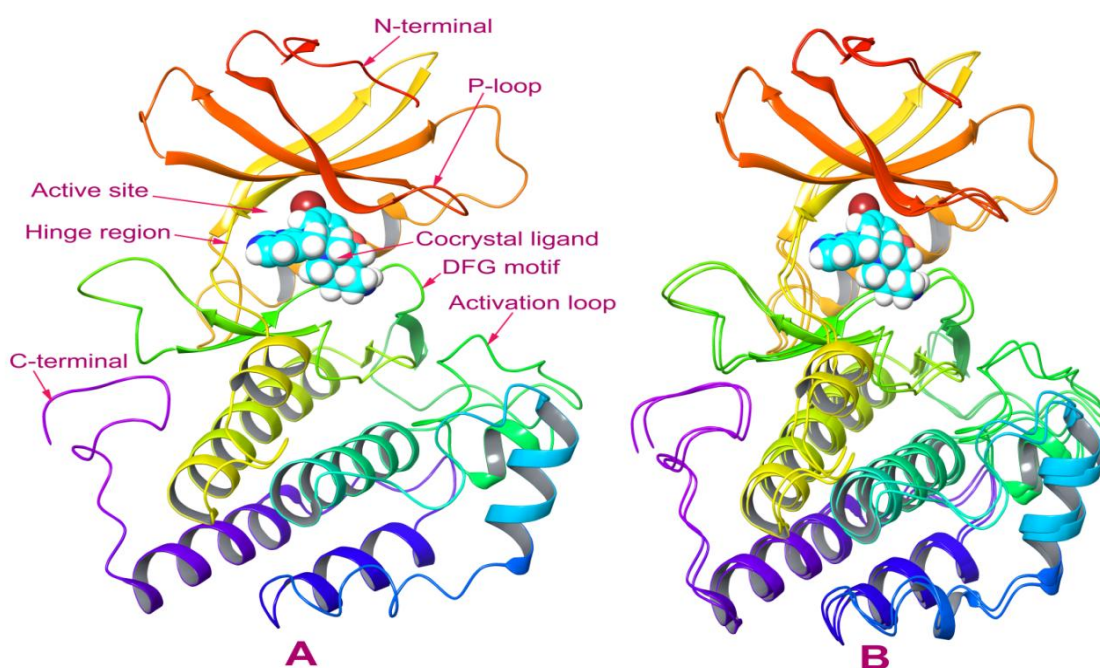
Figure 4.1 Crystal structures of tau-tubulin kinase 1: A, Structure of TTBK1 with the clarification of different regions present in enzyme protein utilizing PDB id: 4NFN; B, Superimposed structures of TTBK1-2KC complex (PDB id: 4NFN), and apoprotein TTBK1 (PDB id: 4NFM).
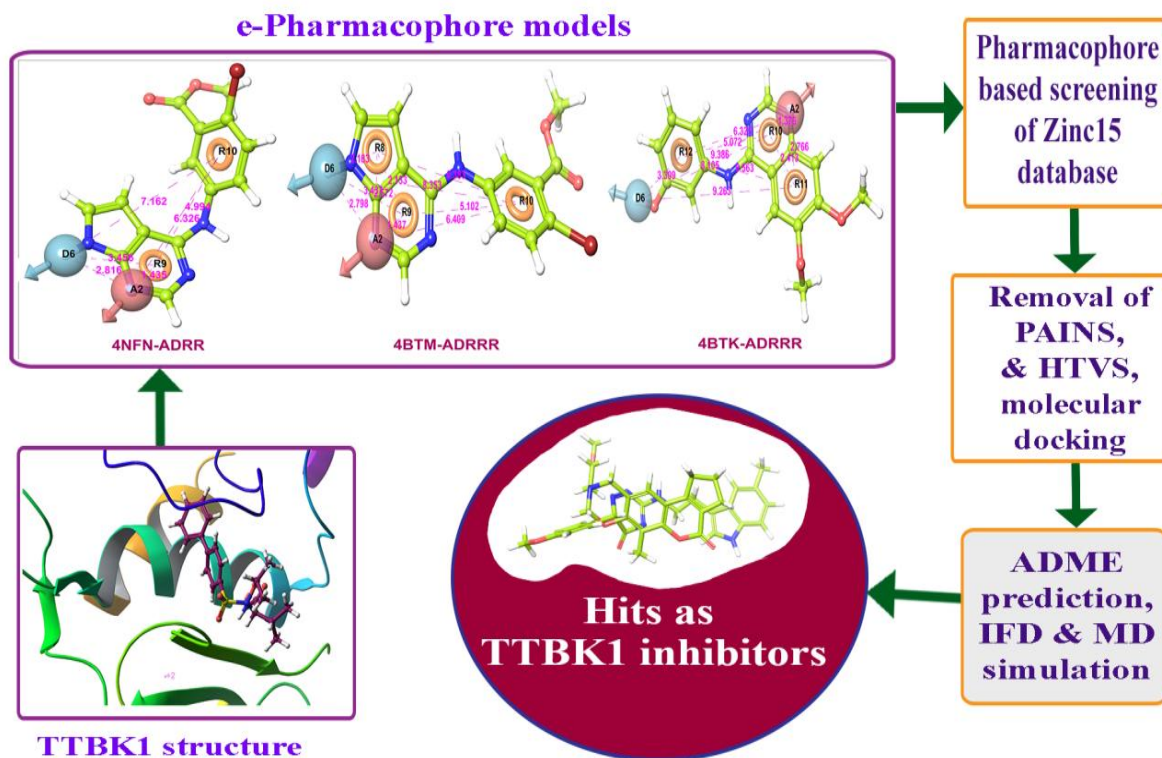

Figure 4.2 Schematic representation of hits identification process as tau-tubulin kinase 1 inhibitor.

4.2 Materials and Methods

4.2.1 Details of computational work tools

The details of computational work tools for development of structure-based pharmacophore models for screening of the 'ZINC15' database and molecular dynamics simulations have been elaborated in Chapter 2, Section 2.2.1.

4.2.2 Development of energy-optimized structure-based pharmacophore

Five crystal structures, *i.e.,* 4NFN, 4NFM, 4BTK, 4BTJ, and 4BTM, for the kinase domain of human TTBK1 (hTTBK1), were obtained from protein data bank (PDB) (https://www.rcsb.org). Three X-ray crystal structures of hTTBK1 with inhibitors, were selected for development of pharmacophore. The cocrystal ligands of three PDB structures *viz.* 3-{5-[(4-amino-4-methyl piperidin-1-yl)methyl]pyrrolo[2,1-f][1,2,4] triazin-4-yl}amino)-5-bromophenol (2KC, IC$_{50}$ 120 nM, PDB id 4NFN, resolution 1.42Å) [12]; methyl 2-bromo-5-(7H-pyrrolo[2,3-d]pyrimidin-4-ylamino)benzoate (F8E, Kd 4.1 μM, PDB id 4BTM, resolution 2.45Å) [13]; and 4-[3-hydroxyanilino]-6,7-dimethoxyquinazoline (DTQ, Kd 240 nM, PDB id 4BTK, resolution 2.0Å) [13], are presented in **Figure 4.3**. Protein structures were prepared by using Protein Preparation Wizard in Maestro 10.1 with an OPLS_2005 force field. The Receptor Grid Generation tool in Maestro 10.1 was applied to prepare Grids of all three protein crystal structures concentrated at cocrystal ligand.
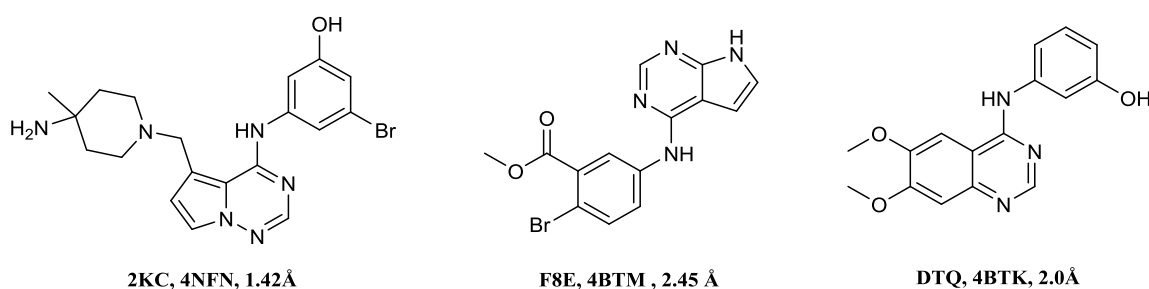

2KC, 4NFN, 1.42Å   **F8E, 4BTM , 2.45 Å**   **DTQ, 4BTK, 2.0Å**

Figure 4.3 Structures of cocrystal ligands with their PDB ids and resolutions.

Cocrystal ligands were refined utilizing LigPrep in Maestro 10.1 with Epik and an OPLS_2005 force field and were docked with 'write XP descriptor information' to their respective prepared protein structures by Glide XP (extra precision) docking. Pharmacophoric sites of the cocrystal ligands were arranged based on their Glide XP docking's binding energies and selected for development of pharmacophore model [5]. The hydrogen bonds, electrostatic rewards, π–π stackings, π–cations, and other interactions were contained within Glide XP descriptors. PHASE v4.2 was applied to generate pharmacophore features *viz.* H-bond donor (D), H-bond acceptor (A), negative ionizable (N), positive ionizable (P), hydrophobic group (H), and an aromatic residue (R) based on XP energy descriptors. H-bond donor (D) and H-bond acceptor (A) of crystal ligands were pictured as vectors, directed to the corresponding H-bond acceptors and H-bond donors at the binding site of protein structures respectively. Most favorable sites were selected for e-pharmacophore hypothesis development by utilizing excluded volume. Initially, maximum ten pharmacophores were set in PHASE v4.2 to develop pharmacophore models. The developed models were validated, gradually reduced, and most effective one was selected for pharmacophore-based virtual screening.

4.2.3 Validation of energy-optimized structure-based pharmacophore

Enrichment factor (EF) and goodness of hit (GH) were calculated (utilizing Equation **2.1** and **2.2** respectively) from a dataset of compounds, prepared from 1000 drug-like decoys (http://www.schrodinger.com/glide_decoy_set). The average molecular weight was 400 D (the "dl-400" dataset), and four known TTBK1 inhibitors were validated for e-pharmacophore hypotheses. LigPrep in Maestro 10.1 with an OPLS_2005 force field, was utilized to prepare ligands database. EF is the fraction of known actives retrieved after a screening of decoy database compounds [6].

4.2.4 Pharmacophore-based screening of database

The 'drug-like' compounds from 'ZINC15' database were collected (without known TTBK1 inhibitors) utilizing Lipinski's filter. Database molecules were prepared using LigPrep with Epik and an OPLS_2005 force field. Compounds, matched with three validated e-pharmacophores, were screened out against prepared 'ZINC15' database molecules using PHASE v4.2 [14]. Compounds with pharmacophore sites' matching score over or equal to 1.5, were selected for further studies. The matching of compounds' conformer on hypothesis based on root mean square deviation (RMSD), site matching, vector alignments, and volume terms, expressed as fitness score [15]. The tolerance of distance was set up to 2.0 Å as a default set of PHASE for perfect matching. PHASE was initialized to rank the pharmacophore matched database in order of the fitness score ranging from 0 to 3. The descending fitness scores of 2000 molecules from highest for each pharmacophore and scores above 1.5, considered as suitable inhibitors, were docked into the binding sites of TTBK1 crystal structure [5].

4.2.5 Removal of pan-assay interference compounds

A list of structural features, known as pan-assay interference compounds (PAINS) which generate many false positives across screening, was reported by Baell and Holloway [16]. Jasial S. *et al.* reported a large-scale analysis of the behavior of PAINS in biological screening assays [17]. 'ZINC15' database molecules within groups (A) anodyne and (B) clean (PAINS-ok) were selected for virtual screening [18]. Other details are depicted in Chapter 2, Section 2.2.8.

4.2.6 High throughput virtual screening and molecular docking

The PAINS free e-pharmacophore matched molecules were screened out by high throughput virtual screening (HTVS), followed by Glide SP (standard precision) and XP (extra precision) docking at the binding sites of respective crystal structures with

Glide, Maestro 10.1. The cocrystal ligand was centralized for grid generation using Grid Generation tools in Glide. Post-docking MM-GBSA (molecular mechanics energies combined with the generalized Born and surface area) minimization was executed to optimize the geometries of retrieves out, and top 10% of retrieves out from each step were selected for the next level. Finally, all the non-peptide retrieves (peptide compounds are orally degradable) were executed to Glide XP molecular docking, utilizing 4NFN crystal structure to estimate the docking score of selected retrieves.

4.2.7 Induced fit docking

A mixed molecular docking and dynamics method, known as induced fit docking (IFD) [19], where the receptor is flexible in the docking study, has been performed. The prepared hits were docked to the rigid protein using Glide, Maestro10.1 with ligand and protein van der Waals (vdW) scaling 0.5 as default setting. Energy minimization was applied on TTBK1 (PDB: 4NFN) crystal structure. The energy minimization of protein structure was performed using OPLS_2005 force-field. The Glide XP was used for the initial docking, and 20 ligand poses were maintained for protein structural refinement. Prime, Schrödinger 2015-1, was utilized to refine residues within 5.0 Å of ligand poses and induced-fit protein-ligand complexes were developed. After refinements of side-chain and backbone [20], ranking of each of the 20 complexes was done by Prime energy. The complex structures within 30 kcal mol^{-1} (ranking) were redocked for the last stage of Glide docking and scoring. Every ligand was docked into each refined low-energy receptor structure, developed in the refinement step. The binding affinity of each complex was estimated in the Docking Score. The higher negative Docking Score is considered as more favorable binding with active site of TTBK1.

4.2.8 ADME properties and blood-brain barrier permeability prediction

The QikProp in Maestro 10.1 [21] was utilized to predict ADME properties and blood-brain barrier (BBB) permeability of the hits. The neutralization of molecules and development of descriptors were not achieved in the normal mode. Therefore, neutralization of all retrieves was crucial before carrying out QikProp. It predicted 44 significant physicochemically and pharmaceutically applicable descriptors, including principle descriptors, physiochemical properties, log P (octanol/water), QP%, log HERG, Caco-2 cell membrane permeability, MDCK cell permeability, skin permeability log Kp and Lipinski's rule of five, which were essential for rational drug design [22, 23]. Predicted brain/blood partition coefficient, QPlogBB was crucial for BBB permeability of compounds.

4.2.9 Molecular dynamics simulation

Molecular dynamics (MD) simulations is a computer simulation method for studying the physical movements of atoms and molecules in a biophysical system. The trajectories of atoms and molecules are resolved by numerically solved Newton's equations of motion for a system of interacting particles, where forces within the particles and their potential energies are often calculated using interatomic potentials or molecular mechanics (MM) force fields [24].

MD simulations of selected four hits and cocrystal ligand were performed using Desmond v4.3 with the OPLS_2005 force field to develop all peptide interactions [25, 26]. The protein-ligand docked complex (.pv file) from XP docking was taken for salvation by applying open TIP3P (transferable intermolecular potential with 3 points) water model in an orthorhombic box of dimension 20×20×20 Å. Protein-ligand complexes had 12 overall negative charges, but 4NFN with crystal ligand had total 13

negative charges and was neutralized by adding Na^+ counter ion for simulation. Others methods details are depicted in Chapter 2, Section 2.2.20.

4.3 Results and Discussion

4.3.1 Development of energy-optimized structure-based pharmacophore

The crystal structures of TTBK1 with cocrystal inhibitors were selected from PDB. Three crystal structures with TTBK1 inhibitors (IC_{50} 120 nM, Kd 240 and 4100 nM) were selected for generating e-pharmacophore hypotheses. Protein preparation wizard, Maestro10.1 was utilized to prepare the proteins with force field OPLS_2005. The RMSD between cocrystallized ligand and docked ligand was 1.24. The amino acid residues' interactions of refined protein structures with cocrystal ligands and docked ligands were similar. Only the distance of H-bonding was slightly changed, and interactions of cocrystal ligands are depicted in **Figure 4.4**. The refined ligands were docked onto the respective prepared protein structures to develop structure-based pharmacophore hypotheses. The mapping of Glide XP energetic terms onto pharmacophore sites generated e-pharmacophore hypotheses were calculated by structural and energy information between protein and ligand. The number of possible pharmacophore sites and number of accepted pharmacophore sites were four for protein structure 4NFN and five for 4BTM. For crystal structure 4BTK, number of possible pharmacophoric sites was six, but number of accepted sites was five. The generated pharmacophore hypotheses for three protein structures were ADRR, ADRRR, and ADRRR, and pharmacophoric feature scores were between −2.17 to −0.33 (**Table 4.1**).

4.3.2 Validation of energy-optimized structure-based pharmacophore

The enrichment factor (EF) and Goodness of hit score (GH) were evaluated to validate the developed e-pharmacophores for further virtual screening. The decoy set consisted

of 1004 molecules in which four were TTBK1 inhibitors. The Güner-Henry scoring method was used to validate the models, and GH values of three pharmacophore models were greater than 0.6 (**Table 4.2**). All the values of parameters ensured that hypotheses were suitable for pharmacophore-based high throughput virtual screening. The distance between the features of e-pharmacophores was between1.376 to 9.263 Å (**Table 4.3**). Three e-pharmacophore hypotheses, selected after validation, are pictured (**Figure 4.5**) with features and their distances.

Table 4.1 e-Pharmacophore hypotheses and feature scores

PDB	No. of possible Site	No. of accepted site	Hypotheses[a]	Pharmacophore features with score
4NFN	4	4	ADRR	A2: −1.29, D6: −1.42, R9: −0.86, R10: −0.64
4BTM	5	5	ADRRR	A2: −2.17, D6: −2.17, R8: −1.13, R9: −0.92, R10: −0.61
4BTK	6	5	ADRRR	A2: −1.42, D6: −0.33, R10: −0.92, R11: −0.88, R12: −1.47

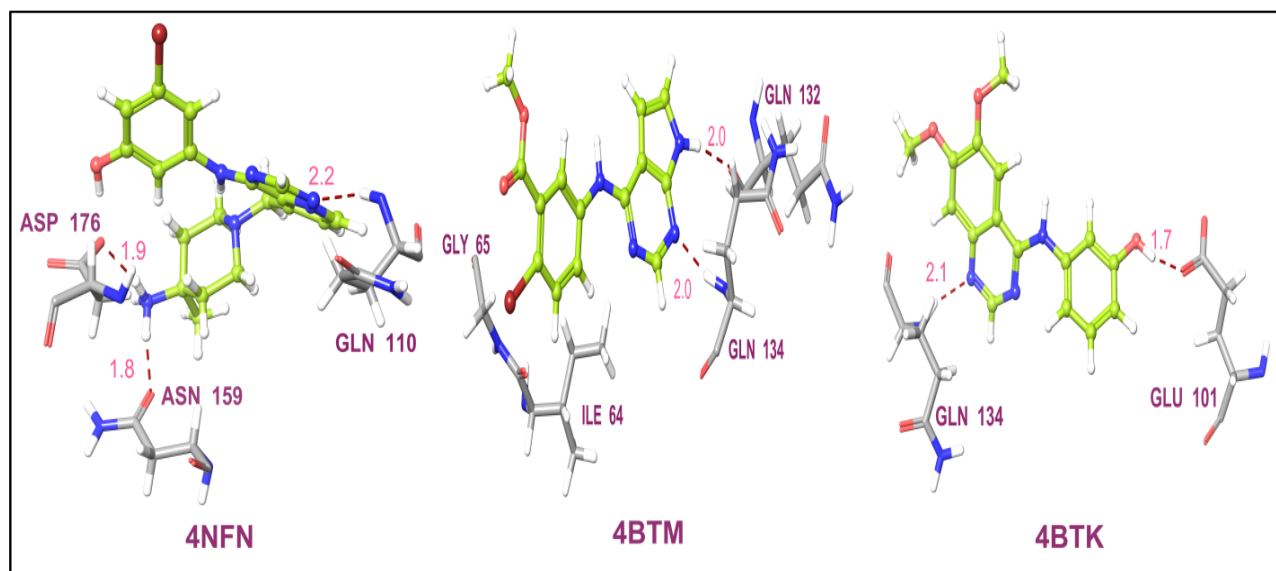[a]A, hydrogen bond acceptor; D, hydrogen bond donor; R, aromatic ring.

Figure 4.4 Residues of crystal structures of tau-tubulin kinase 1 involved with binding interactions of cocrystal ligands (yellowish green).
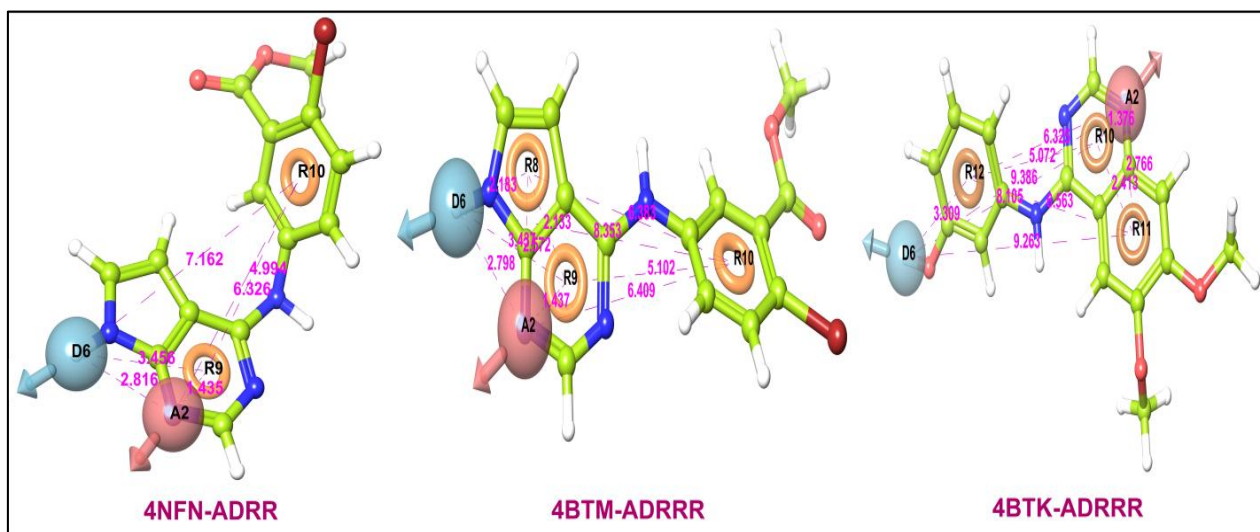
Figure 4.5 e-Pharmacophore models with respective crystal structures. 'A' denotes hydrogen bond acceptor, pink sphere with arrow; 'D' indicates hydrogen bond donor, sky blue sphere containing arrow; 'R' denotes aromatic ring, yellow circle.

Table 4.2 Validation of e-pharmacophores by Güner-Henry scoring method

Parameter	4BTK	4BTM	4NFN
EF[a]	6.3	4.38	9.06
GH[b]	1.63	1.47	1.9
BEDROC(α=8.0)[C]	0.216	0.127	0.282
BEDROC(α=20.0)	0.142	0.066	0.213
BEDROC(α=160.9)	0.116	0.017	0.249
ROC[d]	0.63	0.56	0.64
RIE[e]	2.20	1.03	3.46
AUAC[f]	0.63	0.56	0.63

[a]Overall enrichment factor; [b]Goodness of hit score; [C]Boltzmann-enhanced discrimination of receiver operating characteristic; [d]Receiver operating characteristic; [e]curve value; [f]Robust initial enhancement; [g]Area under accumulation curve.

Table 4.3 Distance between features of e-pharmacophore hypotheses

Pharmacophore model[a]	Distance from A to D (Å)	Distance from A to R (Å)	Distance from D to R (Å)	Distance from R to R (Å)
4NFN-ADRR	2.816	1.435	3.456	4.994
		6.326	7.162	
4BTM-ADRRR	2.798	2.572	2.183	2.133
		1.437	3.437	6.383
		6.409	8.353	5.102
4BTK-ADRRR	9.386	1.376	8.105	2.413
		2.766	9.263	5.072
		6.325	3.309	6.563

[a] PDB used for respective models; A, hydrogen bond acceptor; D, hydrogen bond donor; R, aromatic ring.

4.3.3 High throughput virtual screening and molecular docking

Pharmacophore matched, PAINS free compounds were considered for HTVS to identify hits. Molecular SP and XP dockings were performed by using total 57 HTVS retrieves with 4NFN crystal structure, to compare docking scores. The number of hits from pharmacophore-based virtual screening and final selected hits with their respective PDB are presented in **Table 4.4**. The 22 hits (**Table 4.5**) obtained, after XP docking with 4NFN, had comparatively better docking scores than cocrystal ligand, 2KC. But four hits, *i.e.,* 6-Hydroxyluteolin 7,3'-dimethyl ether; 3-phenyl-2-(9H-purin-6-ylamino) propan-1-ol; 1-[3-(6-aminopurin-9-yl)propyl]-3-methyl-pyridin-2-one; and N-[(4-ethoxy-3-fluoro-phenyl)methyl]-7H-purin-6-amine (ZINC14644839, ZINC00012956, ZINC91332506, and ZINC69775110 respectively) with diverse structures,, PAINS free (except ZINC00012956, a PAINS-ok compound), best docking scores (−10.71 to −8.48 kcal mol^{-1}) and Glide energies (−50.99 to −45.16 kcal mol^{-1}) were selected for studies (**Table 4.5**).

Table 4.4 Compounds retrieved at each stage of screening in dataset

Pharmacophore model	PDB id	PHASE find matches hits	HTVS hits	SP hits	XP hits	Final selected hits
ADRR	4NFN	2000	1942	193	19	ZINC00012956
ADRRR	4BTM	2000	1946	197	19	ZINC91332506 ZINC69775110
ADRRR	4BTK	2000	1929	194	19	ZINC14644839

The number of H-bonding interactions with name and number of interacting residues of protein with hits and cocrystal ligands in ligand-protein is tabulated in **Table 4.6**. The e-pharmacophore of 4NFN afforded one hit, ZINC00012956, and 4BTK yielded ZINC14644839, whereas, e-pharmacophore of 4BTM provided two hits, *i.e.*, ZINC91332506, and ZINC69775110. Chemical structures of selected hits as TTBK1 inhibitor are drawn in **Figure 4.6**. Gln110, the most important residue at the binding site

of TTBK1, produced H-bond interaction with all four hits as well as with 2KC. H-bonding interaction of residue Glu77 with 2KC (**Figure 7A, and C**) and ZINC14644839 are also depicted (**Figure 8A, and C**). Compound ZINC00012956 (**Figure 9A, and C**) interacted through different hydrogen bonding to ASP176, whereas, ZINC91332506 (**Figure 10A, and C**) formed H-bonding with Gly111, and ZINC69775110 (**Figure 11A, and C**) had binding only with Gln110.

Table 4.5 Ligands with better docking score than 2KC after Glide XP docking

Compound	Fitness	Align score	Docking score	Glide energy (Kcal mol^{-1})
ZINC14644839	1.728	0.632	−10.505	−52.819
ZINC89596336	2.073	0.319	−9.772	−46.131
ZINC05999143	1.75	0.619	−9.453	−49.052
ZINC18185774	1.974	0.602	−9.369	−42.645
ZINC17887543	2.078	0.601	−9.368	−43.072
ZINC00012956	1.85	0.712	−9.317	−40.653
ZINC40591228	2.067	0.258	−9.298	−45.85
ZINC75056181	2.012	0.37	−9.293	−37.424
ZINC81004092	1.851	0.653	−9.286	−48.589
ZINC81004171	1.821	0.538	−9.184	−47.723
ZINC09135015	2.164	0.413	−9.08	−42.448
ZINC08553676	1.707	0.716	−8.97	−46.327
ZINC69775110	2.055	0.546	−8.956	−44.481
ZINC69742848	1.868	0.581	−8.878	−41.467
ZINC21992187	1.891	0.622	−8.845	−45.62
ZINC00057845	2.099	0.604	−8.686	−41.676
ZINC91332506	1.963	0.534	−8.533	−43.655
ZINC91332538	1.629	0.732	−8.519	−41.044
ZINC16758238	2.436	0.205	−8.491	−34.957
ZINC47511172	2.036	0.54	−8.466	−39.994
ZINC80981223	1.844	0.411	−8.448	−54.157
ZINC80944903	2.316	0.293	−8.429	−39.414

4.3.4 Induced fit docking

Induced fit docking (IFD) scores of all hits manifested better results than the Glide XP docking scores (**Table 4.6**). The conformations generated from IFD were similar to the

docked poses produced from rigid receptor docking. The Glide-based model provided an RMSD of 4.2 Å for TTBK1, when correlated to the native pose in TTBK1 structure. IFD docking scores of selected hits had better binding energies than 2KC. The number and pattern of H-bond interactions produced after IFD (**Figure 4.7B**, **4.8B**, **4.9B**, **4.10B**, and **4.11B**), were almost similar to Glide XP docking arrangement. Only compounds ZINC91332506, and ZINC69775110 formed additional hydrogen bonding with Gly111 amino acid residue, which also increased overall bonding energies.
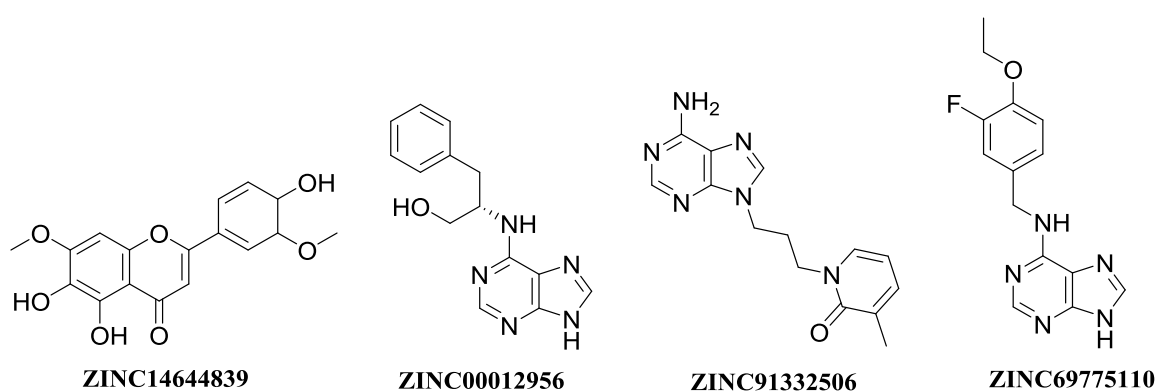

ZINC14644839   ZINC00012956   ZINC91332506   ZINC69775110

Figure 4.6 Structures (with Zinc database ids) of final identified hits as tau-tubulin kinase 1 inhibitors

Table 4.6 Lead molecules and their Glide docking score, number of hydrogen bonds, interacting amino acid residues, and IFD docking score

Compound[a]	Glide docking score	Glide energy	Residue interactions[b]	H-bond	Glide effective state penalty	IFD docking score	IFD score
ZINC14644839	−10.71	−50.99	Glu77, Gln110, Gly111, Asn113	4	0	−11.29	−679.90
ZINC00012956	−8.74	−44.33	Gln110, Asp176	3	0.021	−11.01	−678.94
ZINC91332506	−8.67	−42.48	Gln110, Gly111	2	0	−9.81	−677.49
ZINC69775110	−8.48	−45.16	Gln110	2	0.021	−8.07	−676.89
2KC[C]	−8.37	−43.66	Glu77, Gln110, Asn159, Asp176	3	0.319	−9.44	−678.24

[a]Compounds as per Zinc id, [b]H-Bond interaction, [C]2KC, Cocrystal ligand of 4NFN crystal structure.
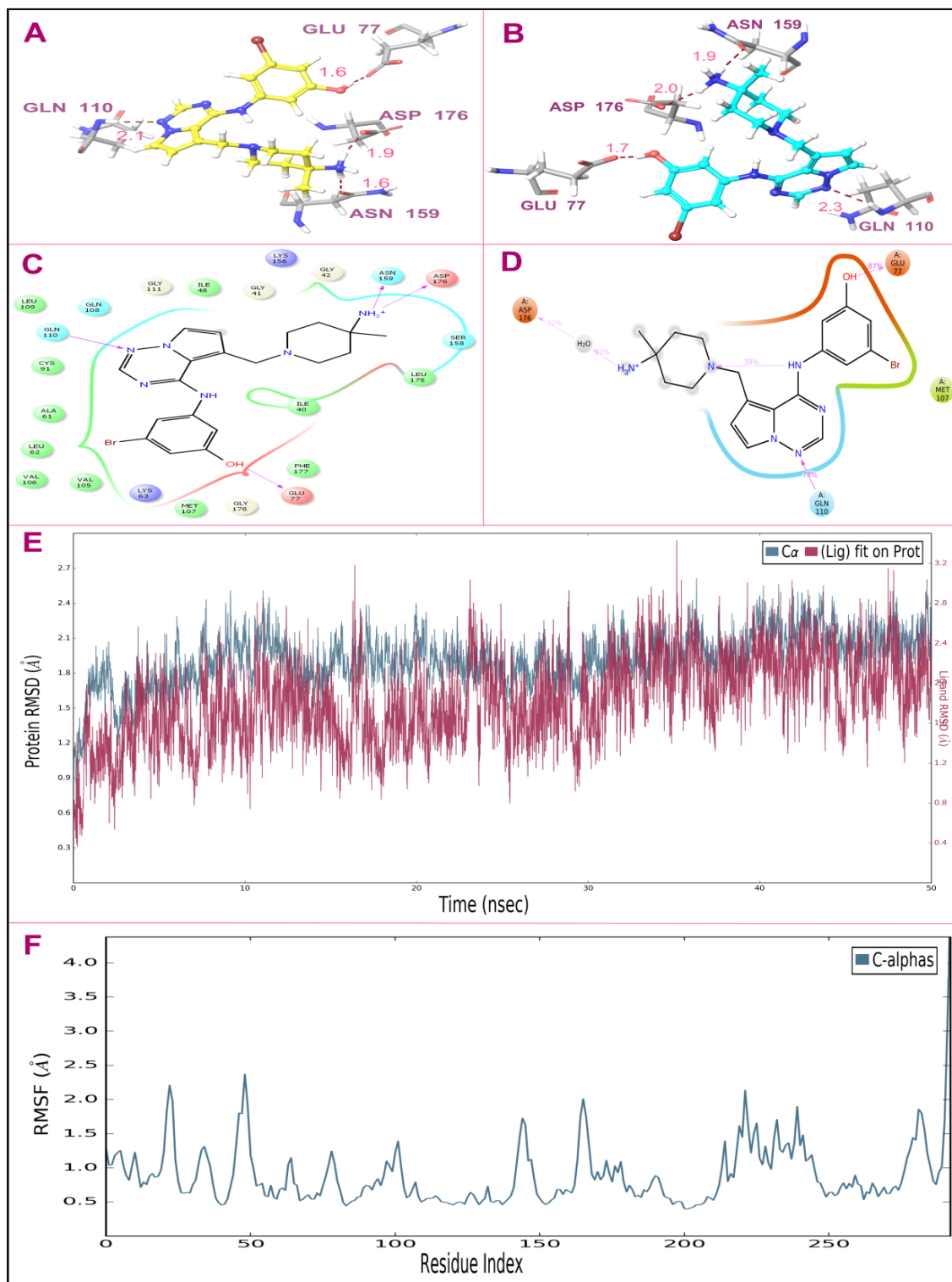
Figure 4.7 Representations for cocrystal ligand 2KC: A, Glide XP docking pose (ligand in yellow); B, Induced fit docking pose (ligand in cyan); C, Glide XP docking 2d pose; D, TTBK1 protein's residue interactions with 2KC after MD simulations, E, RMSD plot of RMSD values for protein on the left y-axis and ligand on right y-axis were displayed, protein backbone in bottle green color, and ligand in maroon; F, RMSF of the protein C-α chain in 2KC-TTBK1 complex.
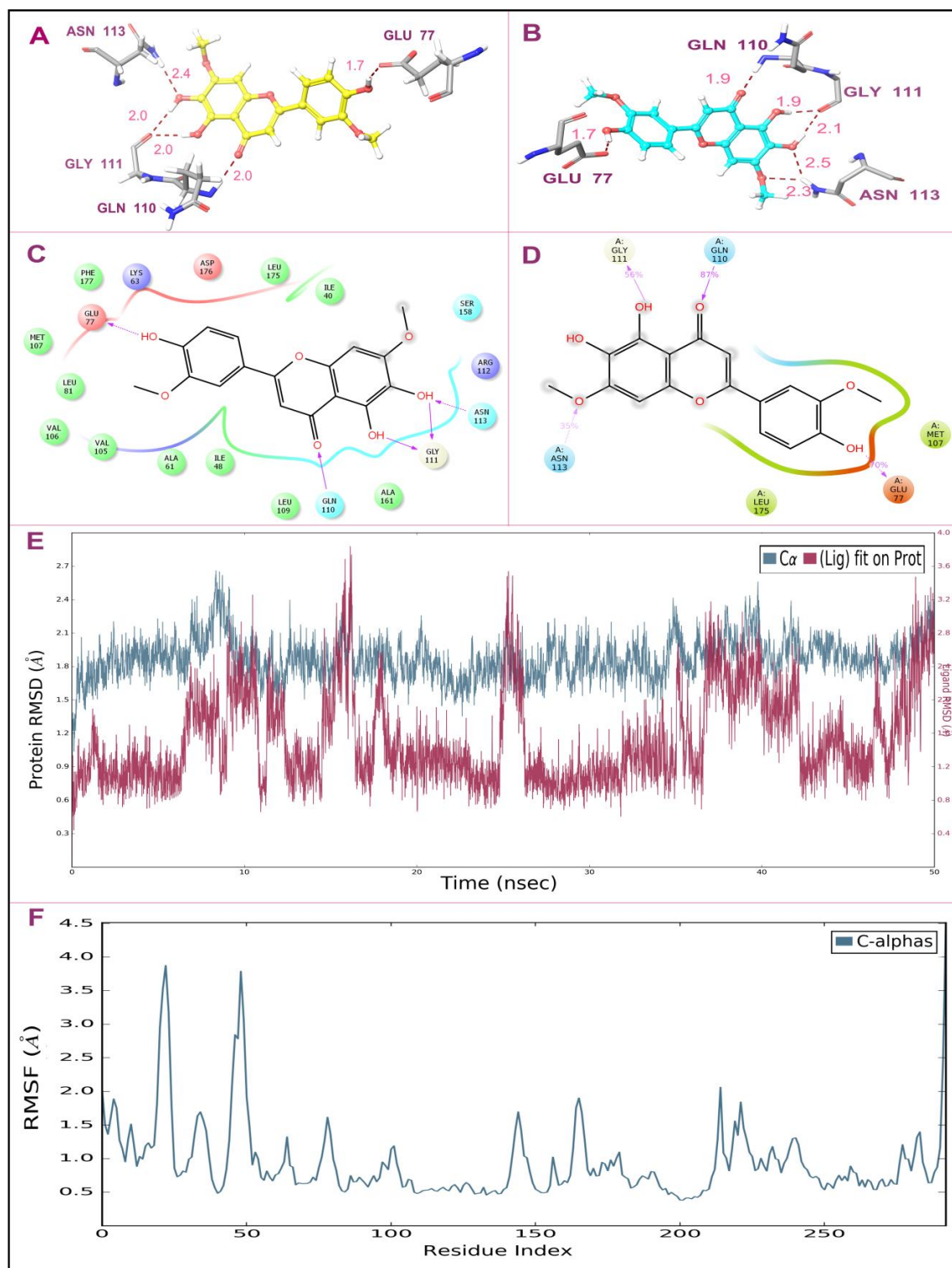
Figure 4.8 Representation for cocrystal ligand ZINC14644839: A, Glide XP docking pose (ligand in yellow); B, Induced Fit docking pose (ligand in cyan); C, Glide XP docking 2d pose; D, TTBK1 protein's residue interactions with ZINC14644839 after MD simulations; E, RMSD plot of RMSD values for protein on the left y-axis and ligand on right y-axis were displayed, protein backbone in bottle green, and ligand in maroon. F, RMSF of the protein C-α chain in ZINC14644839-TTBK1 complex.
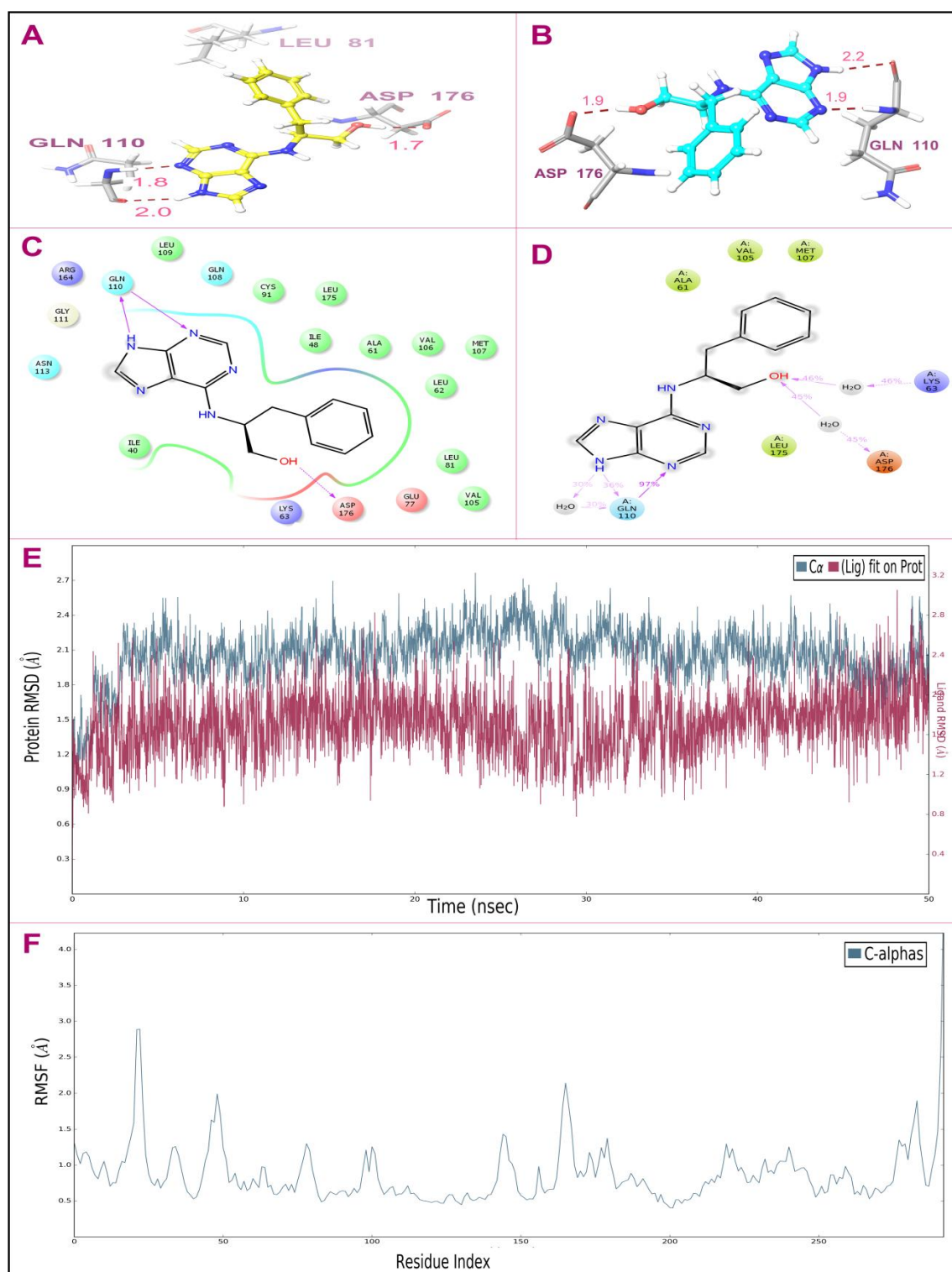
Figure 4.9 Representation for co-crystal ligand ZINC00012956: A, Glide XP docking pose (ligand in yellow); B, Induced Fit docking pose (ligand in cyan); C, Glide XP docking 2d pose; D, TTBK1 protein's residue interactions with ZINC00012956 after MD simulations; E, RMSD plot of RMSD values for protein on the left y-axis and ligand on right y-axis were displayed, protein backbone in bottle green, and ligand in maroon. F, RMSF of the protein C-α chain in ZINC00012956-TTBK1 complex.
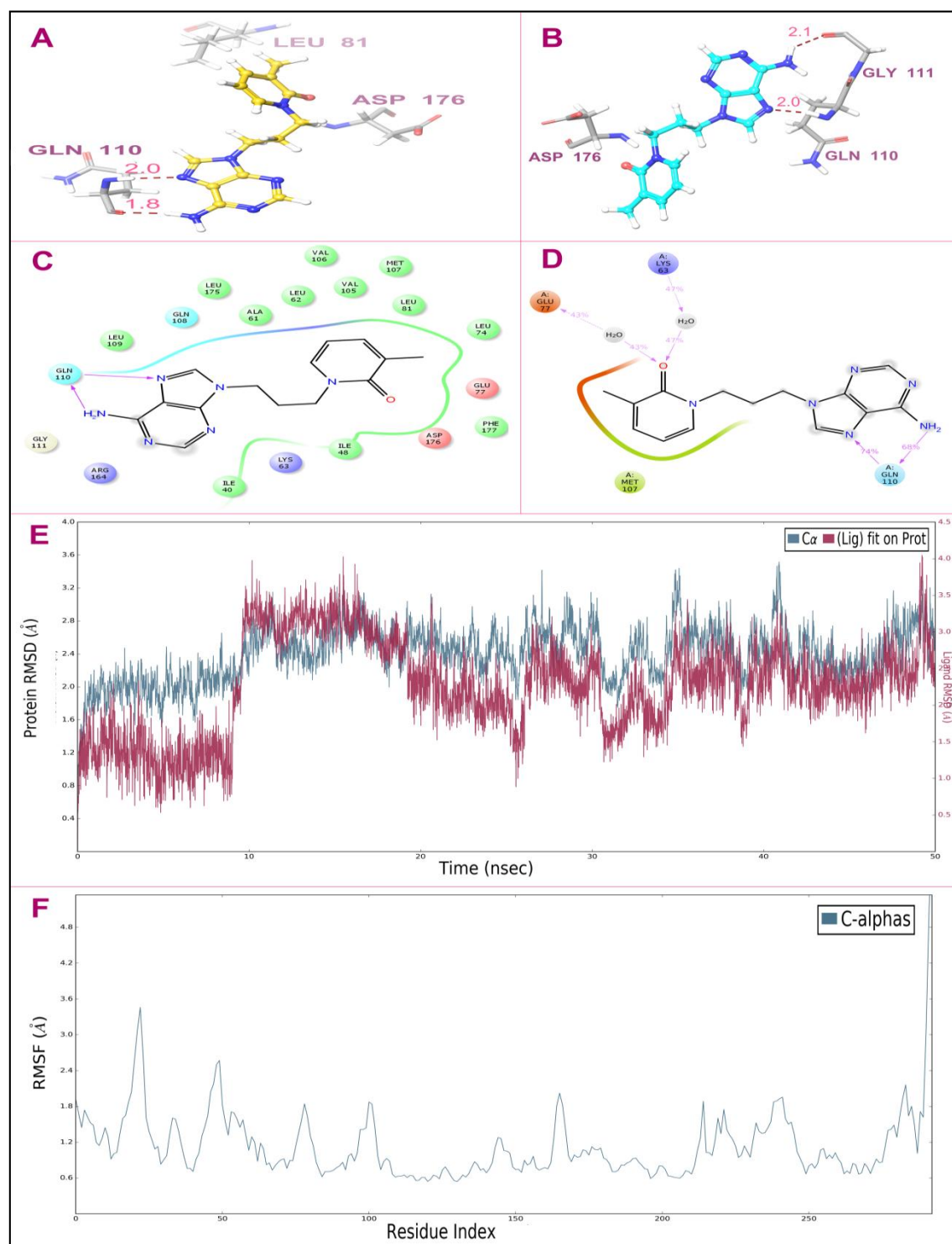
Figure 4.10 Representation for cocrystal ligand ZINC91332506: A, Glide XP docking pose (ligand in yellow); B, Induced Fit docking pose (ligand in cyan); C, Glide XP docking 2d pose; D, TTBK1 protein's residue interactions with ZINC91332506 after MD simulations; E, RMSD plot of RMSD values for protein on the left y-axis and ligand on right y-axis were displayed, protein backbone in bottle green, and ligand in maroon. F, RMSF of the protein C-α chain in ZINC91332506-TTBK1 complex.
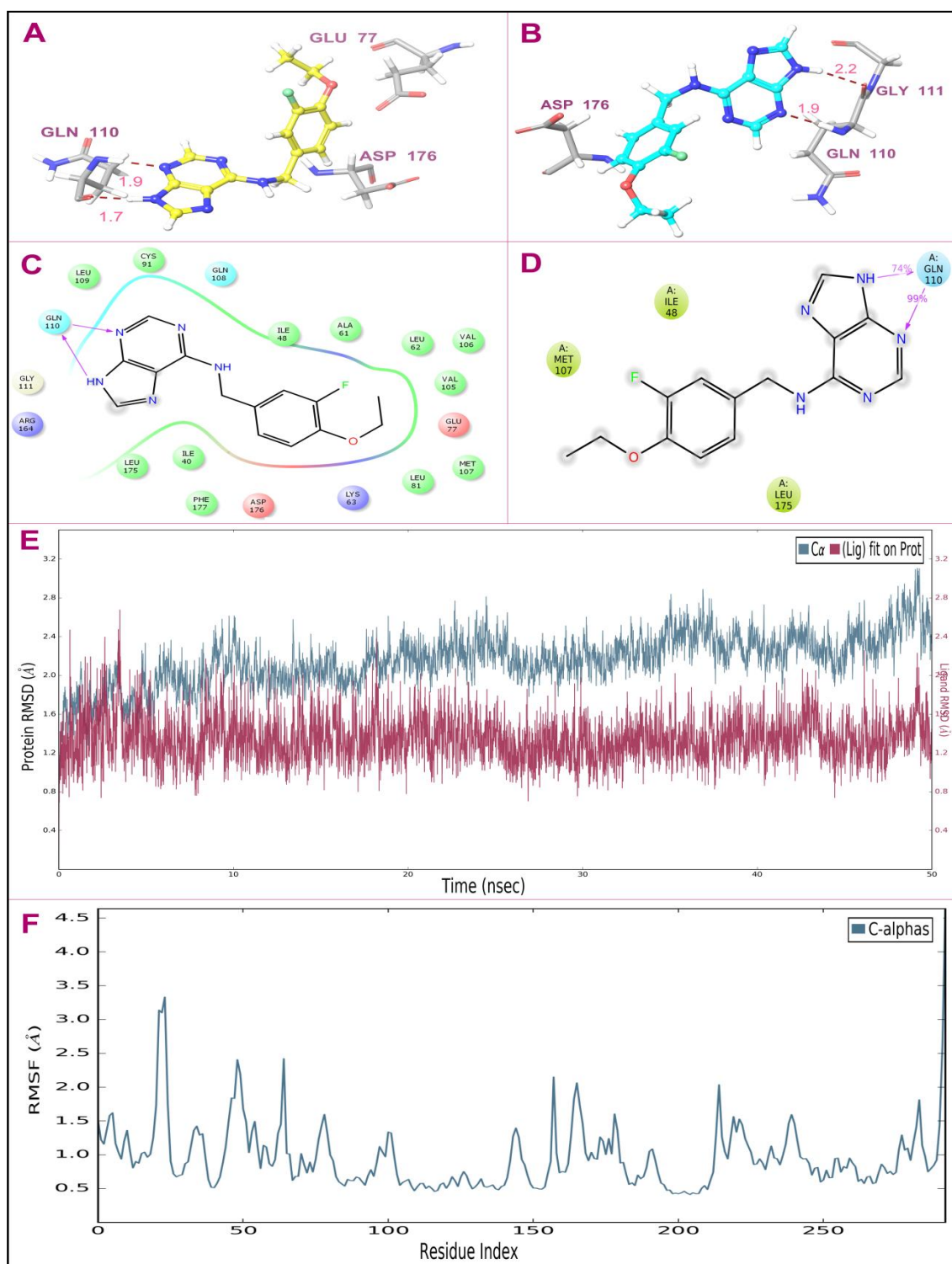
Figure 4.11 Representation for cocrystal ligand ZINC69775110: A, Glide XP docking pose (ligand in yellow); B, Induced Fit docking pose (ligand in cyan); C, Glide XP docking 2d pose; D, TTBK1 protein's residue interactions with ZINC69775110 after MD simulations; E, RMSD plot of RMSD values for protein on the left y-axis and ligand on right y-axis were displayed, protein backbone in bottle green, and ligand in maroon; F, RMSF of the protein C-α chain in ZINC69775110-TTBK1 complex.

4.3.5 Predicted ADME properties and blood-brain barrier permeability

The pharmacokinetics including absorption, distribution, metabolism, and excretion (ADME) were predicted by utilizing QikProp tools of Maestro 10.1. The druggable properties of hits were calculated by analyzing the descriptors physiochemically and pharmaceutically relevant properties (**Table 4.7**). All the hit molecules revealed good partition coefficients (QPlogPo/w) values (1.38 to 2.65), which were necessary for absorption and distribution of drugs inside body. It is an apparent Caco-2 cell permeability in nm s^{-1} value, a factor for the evaluation of cell permeability through biological membranes. The water solubility (QP log S) ranged between −4.09 to −2.93,

Table 4.7 Hit molecules and their physiochemical descriptors determined by Qikprop tools

Compound[a]	QPlog Po/w[b]	QPlog S[c]	QPlogH ERG[d]	QPPC aco[e]	QPPM DCK[f]	QPlog BB[g]	tPSA[h]	QPlog Kp[i]	% Oral Abs.[j]
ZINC14644839	2.07	−3.87	−5.13	206.35	89.85	−1.42	109	−3.58	80.482
ZINC00012956	1.59	−2.93	−5.49	383.26	175.45	−1.18	86	−2.60	82.472
ZINC91332506	1.38	−3.29	−5.31	287.79	128.72	−1.29	92	−3.23	79.018
ZINC69775110	2.65	−4.09	−5.72	997.24	709.84	−0.72	76	−2.07	96.138

Recommended values: [a]Zinc database id and follow the Rule of Five and Rule of Three; [b]QPlog Po/w for octanol/water (−2.0 to 6.5); [c]QPlog S: Predicted aqueous solubility; S in mol dm^{-3} (−6.5 to 0.5); [d] log HERG: HERG K$^+$ channel blockage (<−5); [e] Apparent Caco-2 cell permeability (nm/s) (<25 poor, >500 great); [f] Apparent MDCK permeability (nm/s) (<25 poor, >500 great); [g]QPlogBB: brain/blood partition coefficient (−3.0 to 1.2); [h]tPSA: total polar surface area(Å^2); [i]QPlogKp: skin permeability (−8.0 to −1.0);[j]% Human oral absorption (>80% is high and <25% is poor).

and plogHERG values were less than −5 (−5.13 to −5.72). Hence these could be weaker inhibitors of HERG receptor. The pMDCK (cell-permeable parameter) values were between 89.85 to 709.84, and skin permeability (logKp) values were within −2.07 to −3.58, tPSA values were between 76 to 109Å^2. Predicted brain/blood partition coefficient values of hits were −0.72 to −1.42. BBB permeable parameter values of hits are within accepted range (−3.0 to 1.2) for CNS drugs. The percentage of human oral absorption of molecules was 79.02 to 96.12%, *i.e.,* within the acceptable range for

human use. All the hits possessed entire pharmacokinetic demand for drug-like molecules and followed Lipinski's rule of five as well as rule of three. In conclusion, the hits had better binding interaction than known inhibitor, 2KC and acceptable predicted pharmacokinetic properties and BBB permeability.

4.3.6 Molecular dynamics simulation

The binding modes of cocrystal ligand and protein were obtained from crystallization (PDB); hence the original ligands were more relevant to analyze the identified hits as TTBK1 inhibitors. Although, docking poses developed by Glide, were acceptable and the best poses of ligand-protein interactions were perfect, but were not satisfactory for prediction. A combination of XP docking, IFD and MD simulation of hits were well indicated for stable protein-ligand interactions. The improved docked ligand-protein complexes of all hits by MD, indicated that the selected molecules had a better binding affinity towards TTBK1 than 2KC (**Figure 4.7D**, **4.8D**, **4.9D**, **4.10D**, and 4.**11D**). The pattern and number of interactions of hits (ZINC14644839, ZINC00012956, ZINC91332506, and ZINC69775110) and 2KC with TTBK1 were almost identical from XP docking, IFD, and MD simulation studies. The RMSD of the protein backbone C-α atoms and inhibitor, root mean square fluctuation (RMSF) in the isolated amino acid side chain and ligand-TTBK1 interactions were recorded for a period of 50 ns simulation. The energy of dynamics for ligand-protein complexes was stabilized after 10ns simulation. Robustness and accuracy of MD simulations were indicated by constant temperature, pressure, volume, and potential energy of the complex during the last 40 ns simulation. The RMSD of protein backbone C-α along with the ligand RMSD values, between 0.5 and 3.5Å, indicated the stability of ligand-TTBK1 complexes during simulation. The RMSD plot of RMSD values for protein on the left y-axis and ligand on right y-axis are displayed in **Figures 4.7E**, **4.8E**, **4.9E**, **4.10E**, and **4.11E** with

protein backbone in green, and ligand in maroon. The mean RMSD value for ZINC00012956-TTBK1 (1.05Å), ZINC14644839-TTBK1 (1.25Å), and ZINC69775110-TTBK1 (1.25Å) complexes were lower than 2KC-TTBK1 (1.35Å), whereas the value for ZINC91332506-TTBK1 (1.6 Å) was higher. The RMSF of single amino acid residues throughout the simulation process for all hits and 2KC were below 4.0 Å and were indicative of a lower degree of conformational changes in the side chains. RMSF helped to characterize local changes along the protein chain C-α and peaks denoted the area of protein that fluctuated most during the simulation (**Figures 4.7F**, **4.8F**, **4.9F**, **4.10F**, and **4.11F**).

MD study explored that ZINC14644839 interacted with Glu77, Gln110, Gly111, and Asn113 by direct H-bonding; ZINC00012956 strongly connected Gln110 with direct hydrogen bonding, whereas Lys63, and Asp176 with H-bond formation through a water molecule; ZINC91332506 formed direct hydrogen bonding with Gln110, and H-bonding through a water molecule with Lys63, and Glu77. ZINC69775110 developed direct double hydrogen bonding with Gln110 and, 2KC interacted Glu77 and Gln110 with direct H-bond formation and Asp176 with hydrogen bonding through a water molecule. The interactions of 2KC observed in simulation were fully matched with crystallized structure.MD simulation also supported that the ligands were mostly interacting with protein at Glu77, Gln110, and Asp176 amino acid residues.

4.4 Conclusions

The multiple structure-based pharmacophores integrated with MD simulations afforded ZINC14644839, ZINC00012956, ZINC91332506, and ZINC69775110 as novel TTBK1 inhibitors. The XP docking, IFD and MD simulation studies established that the hits produced stable H-bonding interactions with Gln110, Gly111, Glu77, Asn113, and

Asp176; TTBK1 inhibitor, 2KC, developed hydrogen bonding with Gln110, Asp176, Asn159, and Glu77. The compounds possibly prevent formation of NFTs by inhibiting phosphorylation and aggregation of tau protein and PHFs. Therefore, the hits have prospect to be essential drugs with acceptable ADME properties for the treatment of AD and other neurodegenerative diseases.

4.5 References

[1] Global Burden of Disease Study, Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013, *Lancet (London, England),* 386 (2015) 743–800.

[2] P. Wu, T.E. Nielsen, M.H. Clausen, Small-molecule kinase inhibitors: an analysis of FDA-approved drugs, *Drug Discovery Today*, 21 (2016) 5–10.

[3] K. Iqbal, F. Liu, C.X. Gong, Tau, and neurodegenerative disease: the story so far, *Nature Reviews Neurology*, 12 (2016) 15–27.

[4] K. Iqbal, F. Liu, C.X. Gong, Recent developments with tau-based drug discovery, *Expert Opinion on Drug Discovery*, 13 (2018) 399–410.

[5] S. Sato, R.L. Cerny, J.L. Buescher, T. Ikezu, Tau-tubulin kinase 1 (TTBK1), a neuron-specific tau kinase candidate, is involved in tau phosphorylation and aggregation, *Journal of Neurochemistry*, 98 (2006) 1573–1584.

[6] H. Lund, R.F. Cowburn, E. Gustafsson, K. Strömberg, A. Svensson, L. Dahllund, D. Malinowsky, D. Sunnemark, Tau-Tubulin Kinase 1 Expression, Phosphorylation and Co-Localization with Phospho-Ser422 Tau in the Alzheimer's Disease Brain, *Brain Pathology*, 23 (2013) 378–389.

[7] J. Xu, S. Sato, S. Okuyama, R.J. Swan, M.T. Jacobsen, E. Strunk, T. Ikezu, Tau-tubulin kinase 1 enhances prefibrillar tau aggregation and motor neuron degeneration in P301L FTDP-17 tau-mutant mice, *The FASEB Journal*, 24 (2010) 2904–2915.

[8] S. Sato, J. Xu, S. Okuyama, L.B. Martinez, S.M. Walsh, M.T. Jacobsen, R.J. Swan, J.D. Schlautman, P. Ciborowski, T. Ikezu, Spatial learning impairment, enhanced CDK5/p35 activity, and downregulation of NMDA receptor expression in transgenic mice expressing tau-tubulin kinase 1, *The Journal of Neuroscience*, 28 (2008) 14511–14521.

[9] H. Asai, S. Ikezu, M.E. Woodbury, G.M. Yonemoto, L. Cui, T. Ikezu, Accelerated neurodegeneration and neuroinflammation in transgenic mice expressing P301L tau mutant and tau-tubulin kinase 1, *The American journal of pathology*, 184 (2014) 808–818.

[10] M. Takahashi, K. Tomizawa, K. Sato, A. Ohtake, A. Omori, A novel tau-tubulin kinase from bovine brain, *FEBS Letters*, 372 (1995) 59–64.

[11] M. Goedert, M. Spillantini, R. Jakes, D. Rutherford, R. Crowther, Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease, *Neuron*, 3 (1989) 519–526.

[12] S.E. Kiefer, C. Chang, S.R. Kimura, M. Gao, D. Xie, Y. Zhang, G. Zhang, M.B. Gill, H. Mastalerz, L.A. Thompson, The structure of human tau-tubulin kinase 1 both in the apo form and in complex with an inhibitor, *Acta Crystallographica Section F: Structural Biology Communications*, 70 (2014) 173–181.

[13] Y. Xue, P.T. Wan, P. Hillertz, F. Schweikart, Y. Zhao, L. Wissler, N. Dekker, X-ray Structural Analysis of Tau-Tubulin Kinase 1 and Its Interactions with Small Molecular Inhibitors, *ChemMedChem*, 8 (2013) 1846–1854.

[14] J.J. Irwin, T. Sterling, M.M. Mysinger, E.S. Bolstad, R.G. Coleman, ZINC: a free tool to discover chemistry for biology, *Journal of Chemical Information and Modeling,* 52 (2012) 1757–1768.

[15] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, *Journal of Computer-Aided Molecular Design*, 20 (2006) 647–671.

[16] J.B. Baell, G.A. Holloway, New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays, *Journal of Medicinal Chemistry*, 53 (2010) 2719–2740.

[17] S. Jasial, Y. Hu, J.r. Bajorath, How frequently are pan-assay interference compounds active? Large-scale analysis of screening data reveals diverse activity profiles, low global hit frequency, and many consistently inactive compounds, *Journal of Medicinal Chemistry*, 60 (2017) 3879–3886.

[18] T. Sterling, J.J. Irwin, ZINC15-ligand discovery for everyone, *Journal of Chemical Information and Modeling*, 55 (2015) 2324–2337.

[19] H. Wang, R. Aslanian, V.S. Madison, Induced-fit docking of mometasone furoate and further evidence for glucocorticoid receptor 17α pocket flexibility, *Journal of Molecular Graphics and Modelling,* 27 (2008) 512–521.

[20] S. Saubern, R. Guha, J.B. Baell, KNIME workflow to assess PAINS filters in SMARTS format. Comparison of RDKit and Indigo cheminformatics libraries, *Molecular Informatics*, 30 (2011) 847–850.

[21] E.M. Duffy, W.L. Jorgensen, Prediction of properties from simulations: free energies of solvation in hexadecane, octanol, and water, *Journal of the American Chemical Society*, 122 (2000) 2878–2888.

[22] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Advanced Drug Delivery Reviews*, 23 (1997) 3–25.

[23] F. Ntie-Kang, An *in silico* evaluation of the ADMET profile of the StreptomeDB database, *SpringerPlus*, 2 (2013) 353.

[24] B.J. Alder, T.E. Wainwright, Studies in molecular dynamics. I. General method, *The Journal of Chemical Physics*, 31 (1959) 459–466.

[25] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, *Journal of the American Chemical Society*, 118 (1996) 11225–11236.

[26] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *The Journal of Physical Chemistry B,* 105 (2001) 6474–6487.

[27] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *The Journal of Chemical Physics*, 103 (1995) 8577–8593.

[28] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, *Physical Review A*, 31 (1985) 1695.

[29] G.J. Martyna, D.J. Tobias, M.L. Klein, Constant pressure molecular dynamics algorithms, *The Journal of Chemical Physics*, 101 (1994) 4177–4189.

[30] D.D. Humphreys, R.A. Friesner, B.J. Berne, A multiple-time-step molecular dynamics algorithm for macromolecules, *The Journal of Physical Chemistry*, 98 (1994) 6885–6892.

_____