Journal of Computer-Aided Molecular Design (2022) 36:459–482
https://doi.org/10.1007/s10822-022-00459-0

Deciphering the conformational transitions of LIMK2 active and inactive states to ponder specific druggable states through microsecond scale molecular dynamics simulation

Hemavathy Nagarajan[1] · Ansar Samdani[2] · Vetrivel Umashankar[3] · Jeyaraman Jeyakanthan[1]

Received: 26 February 2022 / Accepted: 16 May 2022 / Published online: 2 June 2022

Abstract

LIMK2 inhibitors are one of the potential therapeutic modalities for treating various diseases. In the current scenario, there is a paucity of effective LIMK inhibitors that are highly specific with minimal off-target effects. To date, the conformational transitions of LIMK2 from $DFG_{in}\alpha C_{in}$ (CIDI) (active) to $DFG_{out}\alpha C_{out}$ (CODO) (inactive) states are yet to be probed and are essential for capturing the unique, druggable conformations. Therefore, this study was intended to capture the diverse conformational states of LIMK2 for accelerating the rational identification of conformation specific inhibitors through high-end structural bioinformatics protocols. Hence, in this study, molecular modelling followed by an extensive microsecond timescale of molecular dynamics simulation was performed encompassing perturbation response scanning, metapath, and community analysis towards the conformational sampling of LIMK2. Overall this study precisely identifies the conformational ensemble of LIMK2 the intermediate inactive states namely, CIDO, CinterDinter, CIDinter, CinterDI, CinterDO, CODI, CODinter apart from CIDI and CODO. This also facilitated observing that β8 preceding XDFG, αC (F373, L374), and αD (L413) as the major effectors that may facilitate the regulation of varying conformational transitions among the states. Additionally, the conserved β sheets and the loops namely, C.l, b.l, and G/P.loop were observed to be involved in the metapath for allosteric communication among the intermediates with CIDI and CODO state. Moreover, only the CODO state was observed to have closed type A.l, while the CIDI and other intermediate states except for CIDO were observed to have open-DFG out type A.l, thereby enabling the binding of substrate. Apart from these, the druggable site analysis inferred that the CIDI and CODO states harbor prominent druggable sites spanning the conserved N-lobe, while the intermediates were observed to have unraveled allosteric druggable sites distal from the ATP binding site, majorly spanning the C-lobe of LIMK2. Thus, this study provides potential insights into the intermediate conformational druggable states of LIMK2 and also the druggable conformations, especially the inactive states of LIMK2, as a specific therapeutic targeting mode. Thus, providing a widened avenue to ponder the allosteric sites or the isoform selectivity conformations for targeting LIMK2 in various disease conditions.

Keywords LIMK2 · Microsecond · DFG · Activation loop · Active · Inactive · Molecular dynamics simulation

Introduction

Protein kinases are one of the major crucial drug targets in humans that are highly dynamic in nature and could adopt varying conformational states to facilitate catalytic activity. It is of high importance to probe and understand the conformational states of protein kinases, as it defines the catalytic activity and protein–protein interactions. In addition, it also assists in determining kinase family-specific inhibitors with higher selectivity and the least off-target effects [1]. In general, protein kinases share a conserved kinase domain comprised of bilobal architecture containing 12 conserved

✉ Jeyaraman Jeyakanthan
jjeyakanthan@alagappauniversity.ac.in

[1] Structural Biology and Bio-Computing Lab, Department of Bioinformatics, Alagappa University, Science Block, Karaikudi, Tamil Nadu 630003, India

[2] Centre for Bioinformatics, Kamalnayan Bajaj Institute for Research in Vision and Ophthalmology, Vision Research Foundation, Sankara Nethralaya, Chennai, Tamil Nadu 600006, India

[3] Department of Bioinformatics, ICMR-National Institute of Traditional Medicine, Nehru Nagar, Belagavi, Karnataka 590010, India

motifs that form secondary structure as follows: 8 α-helices (αB-αI) and 9 β–strands (β1–β9) forming three β-sheets. The bilobal architecture is comprised of an N-terminal lobe (N-lobe) and C-terminal lobe (C-lobe) that are connected by a flexible hinge region, wherein the two lobes form a cleft with a highly conserved ATP-binding pocket. N-Lobe comprises β-strands (β1–β5) and α helices (αB & αC), in addition, the N-lobe contains a conserved ATP-phosphate binding loop known as glycine-rich loop [GxGxϕG] (also called as G/P.loop) that connects β1 and β2 strands, wherein ϕ can be of any hydrophobic residue. The β3 strand comprises a conserved AxK sequence, while glutamate residue is in the middle of the αC helix, are known to be highly conserved. The salt bridge formed between the lysine of the β3 strand and the glutamate of the αC helix is essential for the kinase to be in an "active" state. This conformation is also called αC "IN" (αC_{in}). The absence of this salt bridge denotes the kinase to be "inactive" conformation, where the αC helix attains αC "OUT" (αC_{out}) conformation. The conserved lysine in the β3 strand interacts with α & β phosphate groups of ATP [2].

The C-lobe constitutes the structural elements that are involved in the catalysis namely Asp-Phe-Gly (DFG) motif followed by a flexible activation loop (A.l) [3]. The C-lobe starts with a small helix αD and a long helix αE followed by strand β6 and β-sheets formed by two strands (β6, β9, and β7, β8). The activation loop is located after β9 and is continued by four helices (αF-αI). Among these conserved structural elements, the catalytic kinase domain of protein kinases is defined to be either in its active or inactive state based on the orientation of the highly conserved αC helix and DFG motif of the catalytic domain. The C-lobe contains a catalytic loop (C.l), where the presiding residues are involved in the transfer of the phosphoryl group from ATP to the protein substrates. The catalytic loop that resides near the β7 and β8 strands contains an HRD motif (HRD(x)4 N) which is conserved among the kinases. The aspartic acid (D) in the HRD motif of the catalytic loop places the hydroxyl group of the substrate in position for in-line nucleophilic attack and facilitates the transfer of γ-phosphate of ATP to the substrate. The phosphorylation of the activation loop (A.l) is required for enzyme activity, but not for all kinases. The activation loop is composed of a conserved DFG motif which determines the binding of protein substrate and catalysis of phosphorylation. The conformation of the activation loop is open or extended for the active state enzyme, while the loop remains in closed conformation for the inactive state. The D (Asp) of the DFG motif featuring the activation loop exists in two different conformations. The first conformation is the active state of the enzyme, where the side chain of DFG- "D" points towards the ATP binding site, thereby contributing to the DFG- "IN" (DFG_{in}) conformation. In the inactive state, the DFG-D points to the opposite direction of the binding site, and this conformation is known as DFG- "OUT" (DFG_{out}). In the active cavity, two Mg^{2+} ions coordinate with ATP for the transfer of the phosphate group. The N in the HRD motif (HRD(x)4 N) coordinates with one Mg^{2+} ion, while D in the DFG motif (DFG_{in} conformation) of the activation loop coordinates with the second Mg^{2+} ion [4–7].

Hence, from the varying spatial orientation attained by the highly conserved αC helix and DFG motif of the catalytic domain of protein kinases, a wide range of structural features were derived to define the conformational states of the kinases. The currently available chemical space for kinase inhibitor identification is highly limited, hence there is a need to identify an extended pharmacological space for increasing the rational identification of conformation-specific inhibitors. So far, the major kinase inhibitors are identified to target their active states, yet there are scarce studies on capturing the inactive states and their corresponding inhibitors, thereby identification of these inhibitors may aid in higher specificity. Therefore, understanding the interaction pattern of inhibitors towards conformation specificity has been the current area of interest in kinase research [1, 5, 8–10].

In the current therapeutic research, LIMK inhibitors are emerging as a potential therapeutic modality for treating various diseases like cancer and glaucoma by exploiting the fact that LIMK aids in cell motility and is downstream of Rho-kinase in cytoskeletal dynamics [11]. LIM kinases act downstream of ROCK signaling pathways that regulate the polymerization of actin filaments [12]. Activated LIMK regulates the actin cytoskeleton by phosphorylating its exclusive substrate cofilin/actin-depolymerizing factor (ADF)[13]. This activation mechanism is important for the regulation of cofilin in severing and stabilizing actin filaments. Since phosphorylation of cofilin leads to its inactivation and aids in the destabilization of actin dynamics, it is essential to inhibit the interaction of cofilin with LIMK to regulate the actin cytoskeleton dynamics. Hence, targeting the LIMK will aid in regulating the polymerization of actin filaments and stabilizing the cytoskeleton [13].

Abnormal regulation of LIMK activity has been reported to be involved in the pathogenesis of cancer metastasis, cardiovascular, ocular hypertension, and neurological disorders [14]. Thus, LIMKs are promising therapeutic targets, and specific inhibitors of LIMKs are attractive candidates for novel therapeutic drugs against these diseases. Despite LIMKs being a valid therapeutic target, only a limited number of small molecules are identified for targeting LIMKs however, most of the available inhibitors are not specific and pose toxic side effects, there is a need for an effective LIMK inhibitor that is highly specific with minimal off-target effects. In the current scenario, rational designing of small molecule inhibitors led to the identification of aryl sulphonamides and a pyrrolopyrimidine series of inhibitors with good efficacy, however, none of the

inhibitors have successfully passed the clinical trials. Hence, there is a need to further optimize the existing inhibitors or to identify new compounds with better binding affinity, varied inhibition profiles, and good pharmacokinetic properties for treating various disease conditions [11, 15–17]. To date, the conformational changes of LIMKs that occur from $DFG_{in}\alpha C_{in}$ [CIDI] (active) and $DFG_{out}\alpha C_{out}$ [CODO] (inactive) states of the kinase domain are yet to be probed due to constraints in experimental protocols. Therefore, this study was intended to determine the varying conformational states specifically for LIMK2 to unravel its druggable conformations by implementing high-end structural bioinformatics methods.

In general, computational studies provide quantitative insights into the energetics and kinetics of conformational changes. In specific, various studies have suggested molecular dynamics (MD) as one of the crucial techniques in understanding the metastable states of kinases and determining the conformational changes attained by the kinases. These approaches will support an understanding of the conformational ensembles of the kinases towards rational identification of druggable states and aid in developing conformation-specific allosteric kinase ligands. Hence, in this study, molecular modelling followed by an extensive microsecond timescale of molecular dynamics simulation was performed towards the conformational sampling of the kinase domain. The structural conformations obtained from the MD trajectory (each frame) were extracted and classified into inactive and active states using various descriptors featuring the αC helix and XDFG motif from earlier studies. Wherein, the orientation and positioning of F-[DFG], dihedral of XDFG, pseudo dihedral angle, and salt bridges [αC and β3] were considered for descriptor generation to classify the inactive and active states of DFG and αC helix [4, 18]. Since the inactive states accommodate varying conformations due to the flexibility of the activation loop [A.l] and the G/P loop (G/P.l), thereby the descriptors (A.l and the G/P.l) were incorporated to further classify the inactive states [19]. In addition, the classified conformational states were subjected to perturbation response scanning and community analysis. Further, the classified conformations were probed for pockets that may act as druggable sites for targeting LIMK2. Therefore, this study provides insights into the conformational states of LIMK2 and the conformation-specific druggable sites, especially the inactive states of LIMK2 as a therapeutic target which may aid in identifying the conformation-specific inhibitors for targeting LIMK2 in various disease conditions.

Materials and methods

Refinement and modelling of LIMK2 structures

To determine the conformational transitions of LIMK2, as a first step, molecular modelling and refinement of the available crystal structures were carried out. Only the inactive state of LIMK2 namely, $DFG_{out}\alpha C_{out}$ (CODO) conformation crystal structures are currently available in PDB (PDB ID: 4TPT [20], 5NXD [21]) (Fig.S1). However, these available kinase domain structures had missing residues, E478-N509 following the XDFG and Activation loop (A.l) (Fig.S1). Hence, among the available crystal structures, the selected CODO structure of LIMK2 was further subjected to remodeling and refinement using MODELLER10v [22]. Since the active state [$DFG_{in}\alpha C_{in}$ (CIDI)] of the LIMK2 crystal structure is not yet resolved, the CIDI conformation was modelled using MODELLER10v. As the LIMK2 kinase domain is known to have 70% identity with the kinase domain of LIMK1 [11], the LIMK1 active state (CIDI) structure, PDB ID: 3S95-A was considered as a template for modelling the CIDI conformation of LIMK2. From the generated models, the model with the least DOPE (Discrete Optimized Protein Energy) score was considered the best model and was further validated by the Ramachandran plot. The highly plausible models were further analyzed for their structural stability using molecular dynamics simulation studies.

Molecular dynamics simulation studies

All the Molecular Dynamics simulations were performed using Gromacs2020.1 with Gromos43a1p as the force field, wherein the systems were cubic centered and solvated using the TIP3P water model. Initially, the systems were energy minimized in a vacuum with the steepest descendent algorithm of 50,000 steps. Further, the systems were solvated and neutralized by adding counter ions with two rounds of solvated minimization with the steepest descendent and conjugate gradient of 50,000 steps. The systems were equilibrated with temperature coupling using V-rescale and pressure coupling with Parrinello-Rahman at 300 K and 1 bar. The NVT equilibration of 200 ps was performed with position restraint on the system. NPT equilibration was performed for 100 ps at each step by slowly removing the position restraint by keeping the force constant at 1000 kJ/mol nm^2, 100 kJ/mol nm^2, and 10 kJ/mol nm^2. The bonds were constrained using the LINCS algorithm [23] and Particle-Mesh Ewald (PME) [24] was used for long-range electrostatic calculations. Further, the final MD production run of 1 microsecond (μs) was performed for each system, saving the conformations for each 10 ps. Subsequently, the resulting trajectories were analyzed using GROMACS utilities to probe the conformational changes and fluctuations. Since elasticity is considered one of the intrinsic property properties of proteins, the protein angular dispersion (PADω) which is denoted as one of the quantitative descriptions was determined for each residue using T-Pad analysis [25], the protein angular dispersion (PADω) was calculated to reveal

the backbone transitions by applying directional statistics on the Ramachandran angles and to infer the flexible and rigid residues, as well as the transition states.

Classification of the MDS-derived structures of LIMK2 using DFG and αC classifiers towards active and inactive state identification

The structures (each frame) from the MD trajectory of both 5NXD and modelled LIMK2-$DFG_{in}\alpha C_{in}$ were extracted respectively and were further classified using a custom python script, either as active or inactive states based on the positioning and orientation of the XDFG and αC using various descriptors as shown in Table S1. Initially, the structures were classified as proposed by Roland L. Dunbrack Jr [18]. to identify the DFG active i.e., D469 in "IN" state [DFG_{in} (DI)], "intermediate" [DFG_{inter} (Dinter)], and "OUT" (inactive state) [DFG_{out} (DO)], based on the orientation and location of the F470 side chain ["F"– XDFG] (Table 1). Accordingly, the orientation of F470 was determined based on two distance criteria as follows: D1 = dist [αC-Glu(+4)-Cα, DFG-Phe-Cζ] and D2 = dist [β3-Lys-Cα, DFG-Phe-Cζ], hence the distance of M380 (αC) to F470 (DFG) and K360 (β3) to F470 (DFG) was calculated for DFG classification. Subsequently, these structures were classified further based on the conformation states defined by the backbone dihedral angles (Φ, Ψ) of X-D-F [A468-D469-F470] and the side-chain rotamer of F470 (DFG) (Table 1). The resulting unclassified DFG states from the dunbrack criteria were classified into either DFG_{in} or DFG_{out} states based on the pseudo-dihedral angle (ξ_0)[4], wherein, the ξ_0 of XDFG [A468, D469, F470, G471] and ξ_0 of DFG + 1 [D469, F470, G471, L472] were determined, respectively. Further, all the classified DFG states were considered for determining their respective active and inactive states of αC, based on the criteria proposed by earlier studies [18] as shown in Table 1.

Since the salt bridge formed by K360 (β3) and E376 (αC) define the active and inactive state of αC, the distance between E376 (αC) and K360 (β3) as defined by dunbrack classification was considered for αC classification. While based on mobitz [4], both distances between E376 (αC) to D469 (DFG) and the pseudo-dihedral angle of DFG + 1 (D469, F470, G471, L472) were considered for αC classification. Finally, the classified active, intermediate, and inactive conformations based on DFG and αC orientation were considered for further analysis.

Identifying potential effectors by perturbation response scanning

To investigate mechanisms of allosteric communication and also the series of hot spots that serve as sensors and effectors of allosteric signalling, perturbation-response scanning (PRS) was performed. This technique is considered, as it is crucial for the quantitative measurement of the residues that control the conformational change based on perturbations

Table 1 Descriptors for classifying the active and inactive states of LIMK2

Criteria	LIMK2	Inference
DFG classification by distance		
D1	Cα (380 M)–Cζ (470F)	D1 ≤ 11 Å & D2 ≥ 11 Å → DFG_{in}
D2	Cα (360 K)–Cζ (470F)	D1 > 11 Å & D2 ≤ 14 Å → DFG_{out}
		D1 ≤ 11 Å & D2 ≤ 11 Å → DFG_{inter}
P1	XDFG (468,469,470,471)	< 120° → DFG_{in}
P2	DFG + 1 (469,470,471,472)	> 120° → DFG_{out}
DFG classification by dihedral		
Dihedrals (Φ, Ψ)		Occupancy in Ramachandran plot,
X	A468	"A" for α helical region,
D	D469	"B" for β sheet region,
F	F470	"L" for left-handed α helical region
CHI (χ) angles		
D [χ_2]	D469	χ_1 rotamer of DFG-Phe side chain, "minus" − 60°, "plus" + 60°, "trans" 180°
F [χ_1]	F470	
αC classification		
D3	Cα (376E)–Cα (469D)	D3 < 9 Å → αC_{in}
P2	DFG + 1 (469,470,471,472)	D3 > 10.5 Å → αC_{out}
		9 Å < D3 < 10.5 Å → $\alpha C_{dilated/inter}$
D4	Cβ (360 K)–Cβ (376E)	D4 ≤ 10 Å → αC_{in}
		D4 > 10 Å → αC_{out}

applied to the all-atom model of the protein. In this technique, the perturbations will be introduced as local displacements of selected atoms followed by energy minimization; the response is then measured as the relative ability of a residue to induce displacements of other residues versus its propensity to resist change. [26]. To shed light on residues that play a role in establishing the communication in the kinase domain of the active and intermediate/inactive conformations, the PRS approach was implemented from the PRS module of ProDy [27–29]. PRS calculates the effectiveness and sensitivity of residues in propagating allosteric signals, thereby yielding two groups of residues namely, sensors, and effectors [30]. Effectors are the most effective residues whose perturbation has large effects on the structure and dynamics in response to external perturbations. On the contrary, sensors are the most sensitive residues, high propensity to sense signals most strongly to perturbations of effectors.

Protein-structure network and global meta path analysis

Graph theory has been widely used to describe the dynamics and structural properties of proteins. The weak intramolecular interactions in a protein can be collectively represented in the form of a network, i.e. a Protein Structure Network (PSN), where the residues are the nodes of a network and are connected by links (edges) that depend on their interaction strength or their energetic coupling. The residues can communicate through the shortest paths available and multiple paths have been observed between the same pair of distal sites with many nodes in common. The capability of PSN methods to define paths of communication over long distances makes it a suitable technique for the prediction of allosteric effects in protein structural ensembles. PSN approaches can be used to provide a subset of paths of communication between distal residues that are likely to be paths that allow the propagation of structural communication or even predict allosteric hot spots in a protein structure. PSN analysis is a powerful tool to unveil the intra-structural communication in a biomolecular system [31–35].

In this study, a mixed PSN-ENM approach was implemented to predict structural-communication pathways and to determine the shortest communication path in all the classified active and intermediate/inactive conformations using WebPSN [36]. Since comparison of network features is a powerful tool to decipher the structural signatures of functionally different states in different sets of homologous proteins at varying degrees of similarity. In addition, the PSN difference of the intermediate conformations with respect to active (CIDI) and inactive (CODO) states were also predicted. Whereas, the shortest communication pathways were defined by retaining only the shortest path(s) that contains at least one residue correlated (minimum cross-correlation value of 0.6) with either one of the two. In a structural network, selected properties define the calculation of the shortest path in structural regions of protein interaction, and an increase or decrease in these values is related to an increase or decrease in intramolecular interactions. Correlation values above 0.60 represent complexes that contain at least one correlated residue, and this usually allows only the shortest paths to be retained. The strength of the interactions, as well as the number of nodes and predicted links in the trajectories, show that both homologs are capable of altering the conformation and stability of the free protein. Moreover, metapaths which provide an overall representation of structural communication based on most recurrent nodes and links in the path were also calculated. As a comparative analysis of intermediates along with active and inactive states for filtered metapath, only αCE (αC-E376) and DFG (D469) were considered, as these nodes define the conformational states of kinase, whereas minimum frequency was set as 50, Minimum Correlation of 0.7, and Minimum Recurrence as 10. All filtered metapaths were visualized with VMD.

Classification of inactive states of LIMK2

The relative stability of the active and inactive kinase conformations determines the overall activity of kinases. The factors modulating DFG conformations remain to be a mystery. However, several aspects suggest that at least most kinases can adapt every major conformation at a conformational penalty. Energetic accessibility of the different states might be shifted by multiple factors, such as the phosphorylation state of the A.l, binding of allosteric regulators, or induction by inhibitors. Overall, there are a plethora of inactive conformations with varying A.l, DFG, and αC-helix orientations. Apart from DFG and αC, the other segments that influence the variable states of inactive conformations are A.l and G/P.l. The C-terminal domain consists of a flexible A.l, typically 20–30 amino acids in length, and marked by a conserved Asp-Phe-Gly (DFG) motif at the beginning of the A.l. The A.l can undergo large conformational changes, thereby controlling the catalytic activity and access to the substrate-binding pocket [37]. The glycine-rich G/P.loop (G/P.l) is a part of the N-terminal β-sheet region located between strands β1 and β2 which controls ATP binding and phosphate transfer in protein kinases. The three glycine residues in the G/P.l (residues that impose minimal steric interference) make it one of the most flexible elements in the catalytic core [38]. This in turn leads P.l to stack down onto the ligand, creating a more buried cavity and stabilizing interactions with the ligand. Hence, this emphasizes the importance of defining the flexibility of the inactive state of kinases ("DFG_{out}") based on A.l and G/P.l. Therefore, the conformational space accessed by the G/P.l and A.l in

classified intermediate inactive conformations of the LIMK2 was determined based on the A.l and P.l based pseudotorsional angles (ξ_0), dihedrals (Φ, Ψ), and distance as mentioned in Table S1 and S2 [37] using custom python script.

Identification of druggable site and Pocketome comparison analysis

The classified active and inactive conformations were subjected to druggable site analysis and allosteric cryptic site predictions using ESSA [39, 40]. From the resulting cryptic sites, only the sites with a druggability score of > 0.5 and the highest pocketome volume were considered the druggable sites. Among these sites, those with optimal druggable pockets were considered for comparative analysis against allosteric and orthosteric ligands available in the KLIFS database accessed on 24 Nov 2021[3, 41, 42]. For this comparative analysis, all the 12,619 monomers available from KLIFS were considered and aligned to the predicted pockets spanning the classified conformations. A customized python script was used to map the identified druggable pockets of the conformations that overlap with the available allosteric and orthosteric ligands that are feasible for inhibitor identification in further studies. Further, the predicted orthosteric or allosteric ligands that were mapped to the druggable sites in the classified conformations were analyzed against the PKIDB [43, 44], to determine the inhibitors that are in clinical trials for kinases, thereby the significance of the predicted druggable pockets ascertained.

Results and discussion

Structure selection of LIMK 2 and refinement of the kinase structures

Among the crystal structures reported, apart from variation in dimerization, each chain of the LIMK2 dimer structures was observed to have altered variation in G/P.loop angle, distance, and rotation. On structural superimposition of the monomers of respective dimers (Fig. S2), a drastic shift was observed in the G/P.loop (Fig. S2). Furthermore, on analyzing the chi angle of the crucial residues that moderate the catalytic cleft, it was observed that activation loop residue S473 in 4TPT-B, 5NXD-B, and F470 (DFG) only in 5NXD-A varied drastically when compared to other respective monomers (Fig. S2). On considering the gatekeeper residue (T405) and hinge domain residues (E406-I408), the side chain variation was observed only in 4TPT-A. While 5NXD-A was found to be highly varied than other monomers in terms of all the hinge domain residues. However, in the case of Y407, even 4TPT-B showed significant variation. In addition, the crucial salt bridge forming residues E376 and K360 also showed variation (Fig. S2). Moreover, the 4TPT lacks a proper G/P.loop, as it precedes the loop conformation, whilst 5NXD showed an extended G/P.loop followed by β-sheets; in addition, it was observed that 5NXD had the least distance between the αC and DFG motif. Hence, 5NXD-A (LIMK2 $DFG_{out}\alpha C_{out}$ state) was selected for modeling the missing residues region extending the XDFG and activation loop (A.l). The best refined model of 5NXD-A (least DOPE score) (Fig. 1a) showed no residues in the disallowed region of the Ramachandran plot and proceeded for further studies.

In the case of active state conformation of LIMK2 kinase domain modelled based on LIMK1(PDB ID:3S95-A[$DFG_{in}\alpha C_{in}$]), the best model (least DOPE score) (Fig. 1b) was observed to have residues with no outliers in the Ramachandran plot analysis. On structural superimposition of the modelled $DFG_{in}\alpha C_{in}$ with $DFG_{out}\alpha C_{out}$ (Fig. 1c), it was observed that "F470" of the DFG motif was found to be in alternate orientation, i.e., the "F470" was pointing inwards in the inactive state i.e., $DFG_{out}\alpha C_{out}$ facilitating the hydrophobic pocket, thereby hindering the interaction of ATP. Whereas, in the active state ($DFG_{in}\alpha C_{in}$), the "F470" was protruding outwards, thus facilitating the "D469" of the DFG motif for feasible interaction with the ATP to exhibit its catalytic activity. Moreover, the conformation state of the modelled $DFG_{in}\alpha C_{in}$ structure was also verified based on the Chi and dihedral angles of the DFG motif and αC as shown in Fig. S3, whereas the DFG and crucial αC residues of the $DFG_{in}\alpha C_{in}$ state were found to be in altered position than the inactive state. This plausible structure of 5NXD-A ($DFG_{out}\alpha C_{out}$ state) and modelled $DFG_{in}\alpha C_{in}$ of LIMK2 were considered for further molecular dynamics simulation studies.

Molecular dynamics simulation

From the microsecond scale MD trajectory of the $DFG_{in}\alpha C_{in}$ and $DFG_{out}\alpha C_{out}$ (5NXD-A) states of LIMK2, it was observed that the $DFG_{out}\alpha C_{out}$ (inactive) state was found to have a prominent backbone RMSD (Root Mean Square Deviation) of about ~ 0.6 nm than the $DFG_{in}\alpha C_{in}$ (active) state (~ 0.38 nm) throughout the MD simulation (Fig. 2a). However, both the states were observed to attain equilibrium after 600 ns of the MD production run. On analyzing the comparative RMSF (Root Mean Square Fluctuation) plot of the $DFG_{in}\alpha C_{in}$ and $DFG_{out}\alpha C_{out}$ (5NXD-A) states of LIMK2 (Fig. 2b and c), overall C-lobe residues were observed to be highly fluctuating than the N-lobe residues. Among the Loop conferring residues, in the case of $DFG_{in}\alpha C_{in}$, K339 (G/P-loop), M380 (regulatory spine-αC), catalytic spine residue (L413 (αD), I459 (β7)), Y407 (hinge), the D398 that connects β4 and β5 were found to be highly fluctuating. While in

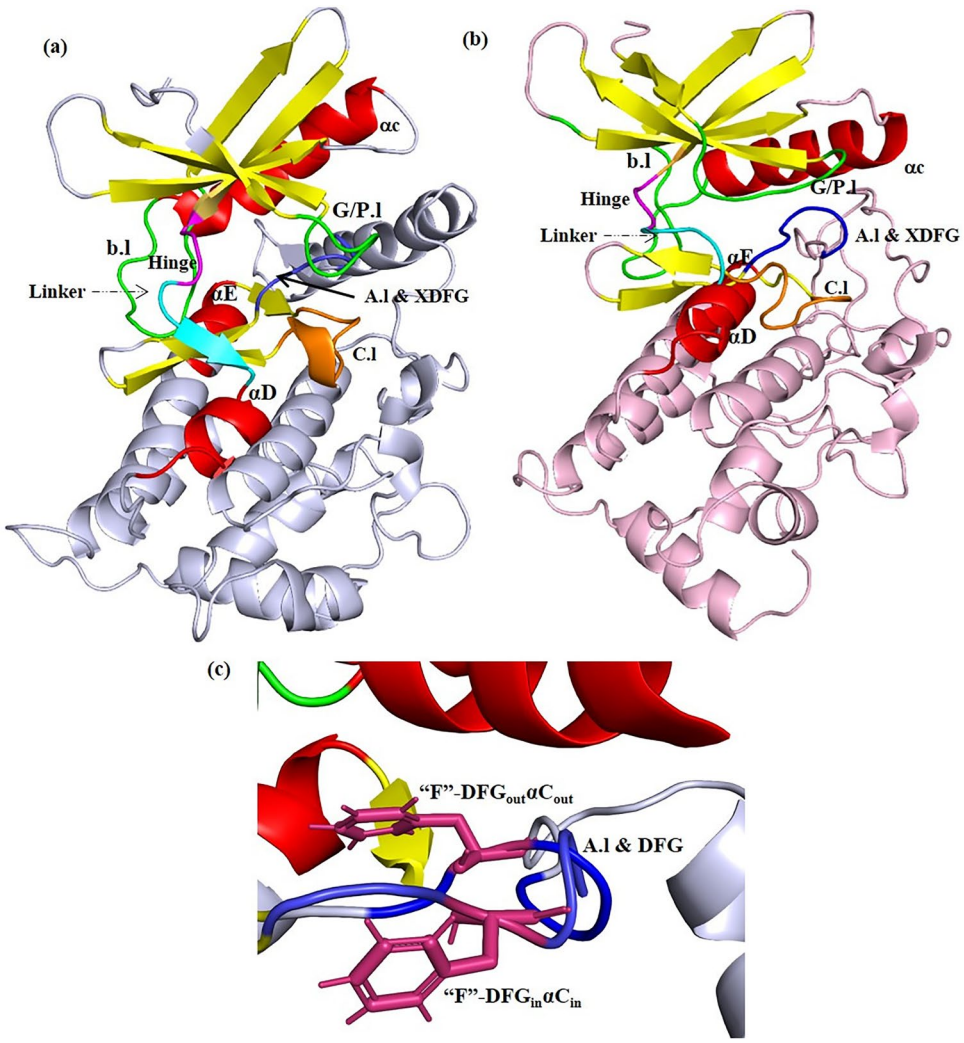

Fig. 1 Modelled and refined structure of LIMK2 (**a**) inactive state [$DFG_{out}\alpha C_{out}$]– 5NXD-A (**b**) active state [$DFG_{in}\alpha C_{in}$]; where the conserved kinase structural elements are colour coded with β sheets are coloured in yellow, the loops namely b.l (back loop), C.l (Catalytic loop), A.l (Activation loop) are shown in green colour (**c**) variation in the placement of "F"-DFG in the active and inactive states, where "F" –DFG is shown as sticks and pink coloured

$DFG_{out}\alpha C_{out}$ only the K390, F391 of the back loop (b.l) that connects αC and β4, and L472 (A.l) were found to fluctuate.

Classification of the active and inactive states of kinases

The MD trajectory structures extracted per frame (~ 1,00,000) from each $DFG_{in}\alpha C_{in}$ and $DFG_{out}\alpha C_{out}$ (5NXD-A) trajectories were extracted and considered for conformational classification of LIMK2. From the distance-based DFG classification of $DFG_{out}\alpha C_{out}$ (5NXD-A) trajectory structures, except for 16 structures (0.02%), all other structures (99.98%) were found to have DFG_{out} conformation (Fig. 3a). However, the unclassified structures (0.016%) on pseudotorsion angle classification resulted in DFG_{out} conformation (Fig. 3b). Thereby, all the structures extracted from $DFG_{out}\alpha C_{out}$ (5NXD-A) trajectory resulted in having DFG_{out} conformation, however, among these structures, only 33.048% were found to be BBA minus class, as per dihedral (Φ, Ψ) and Chi (χ) angle based classification of XDFG motif (Fig. 3c and d). On subsequent αC classification of these DFG_{out} structures, it was inferred that most of the structures nearly 99.72% were observed to have αC_{out} conformation, followed by 0.27% with αC_{in} conformation, while only negligible structures (0.01%) were to be in αC inter/dilated conformation (Fig. 3e and f). Therefore, the $DFG_{out}\alpha C_{out}$ (5NXD-A) trajectory resulted in three intermediate/inactive states namely $DFG_{out}\alpha C_{in}$ [CIDO]; $DFG_{out}\alpha C_{inter}$ [CinterDO]; $DFG_{out}\alpha C_{out}$ [CODO]. The DFG and αC orientation of these classified conformations (Fig. S4) showed that E376 of αC protruded out in both CODO and CinterDO, while these were found in close proximity in the case of CIDO conformation. Despite the inward projection of F470 (DFG) in all these conformations, it was observed that the orientation of F470 towards the ATP cavity was different in CIDO when compared to the other two conformations (CODO and CinterDO). From the T-PAD analysis (Table S3), it was observed that XDFG residues feature short (38.3°—51.61°) and long transitions (32.56°–41.33°). In the case of CinterDO, except for D469, other residues

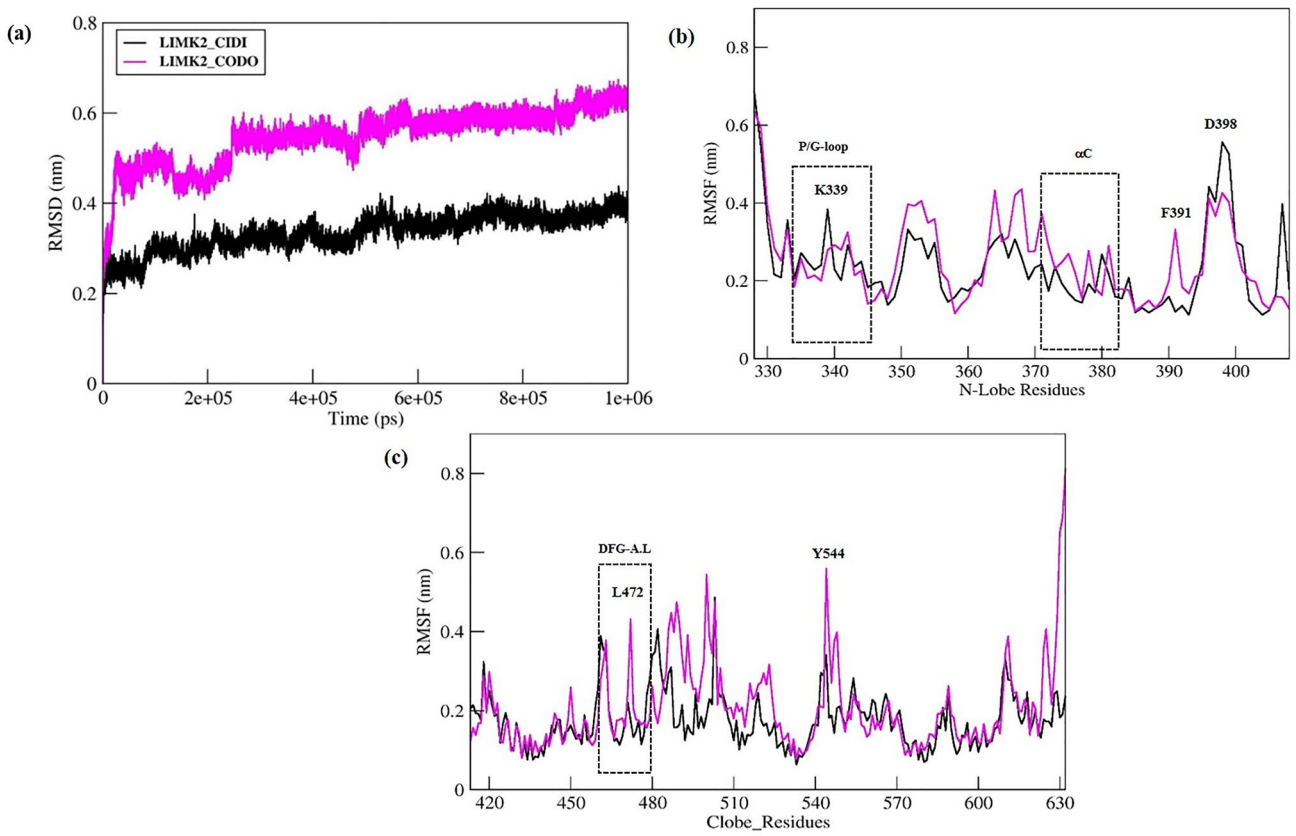

Fig. 2 Molecular dynamics simulation analysis of LIMK2 –CIDI (Active) and CODO (inactive) states (**a**) RMSD; Residue fluctuations RMSF plot (**b**) N-lobe (**c**) C-lobe

were found to have long transitions (86.9°- 28.14°) in CIDO conformation and fluctuations (48.85°–39.01°) in CODO. Whereas D469 showed higher fluctuations in CIDO and long transitions in CODO. The F470 residue of DFG that defines the catalytic state of kinase and acts as a well-defined regulatory spine was observed to have undergone a higher PAD degree of long transition (81.04°) in CIDO and fluctuation (41.31°) in CODO. In the case of αC conformation outlining elements, the salt bridge forming residues, K360 (β3) and E376 (αC) were observed to have fluctuations in both CinterDO and CODO with PAD degrees of about 36.76°, 30.9°, and 45.52°, 25.67°, respectively, whilst CIDO showed a shorter transition. Apart from the αC and DFG, all the catalytic and regulatory spine residues were observed to have fluctuations in CODO conformation. On contrary, in the CinterDO conformation, these residues were found to have long transitions. In specific, the catalytic spine residues A345 and V358 were found to have short transitions in CIDO conformation, while A345 has long transitions in CinterDO. I459 was observed to have long and short transitions in CinterDO and CIDO, respectively.

In the case of $DFG_{in}\alpha C_{in}$ [active] MD trajectory classification, the initial distance-based DFG classification resulted in DFG_{in} (88.26%) and DFG_{inter} (11.74%) states (Fig. 4a). Whereas, the dihedral classification of the DFG_{in} state structures was observed to have various dihedral classes namely, BLAminus (9.38%), BLAplus(0.005%), and ABAminus (1.83%) class. However, the DFG_{inter} structures were found to be only in the BAB trans (10.839%) class (Fig. 4b and c). On the further classification of DFG_{in} conformations for its respective αC state, it was found that DFG_{in} conformation αC accommodates varying orientations thereby resulting in αC_{in} (62.35%), αC_{inter} (24.35%), and αC_{out} (1.77%) (Fig. 5a and b). While the DFG_{inter} conformations were observed to have αC_{in} (98.84%), αC_{inter} (0.67%), and αC_{out} (0.477%) states (Fig. 5c and d). Thus, the classification of the LIMK2 ($DFG_{in}\alpha C_{in}$) trajectory resulted in $DFG_{in}\alpha C_{in}$ [CIDI] (active state), Intermediate states namely $DFG_{in}\alpha C_{inter}$ [CinterDI], $DFG_{in}\alpha C_{out}$ [CODI], $DFG_{inter}\alpha C_{in}$ [CIDinter], $DFG_{inter}\alpha C_{inter}$ [CinterDinter] and $DFG_{inter}\alpha C_{out}$ [CODinter]. On observing the intermediates with the classified CIDI (Fig. S5), it showed both K360 (®3) and E376 (αC) in close proximity favoring salt bridge formation, wherein the F470 "DFG" was found to be displaced away from the ATP binding site. Similarly, the intermediates CODI and CinterDI, and CinterDinter conformations were observed to have the F470

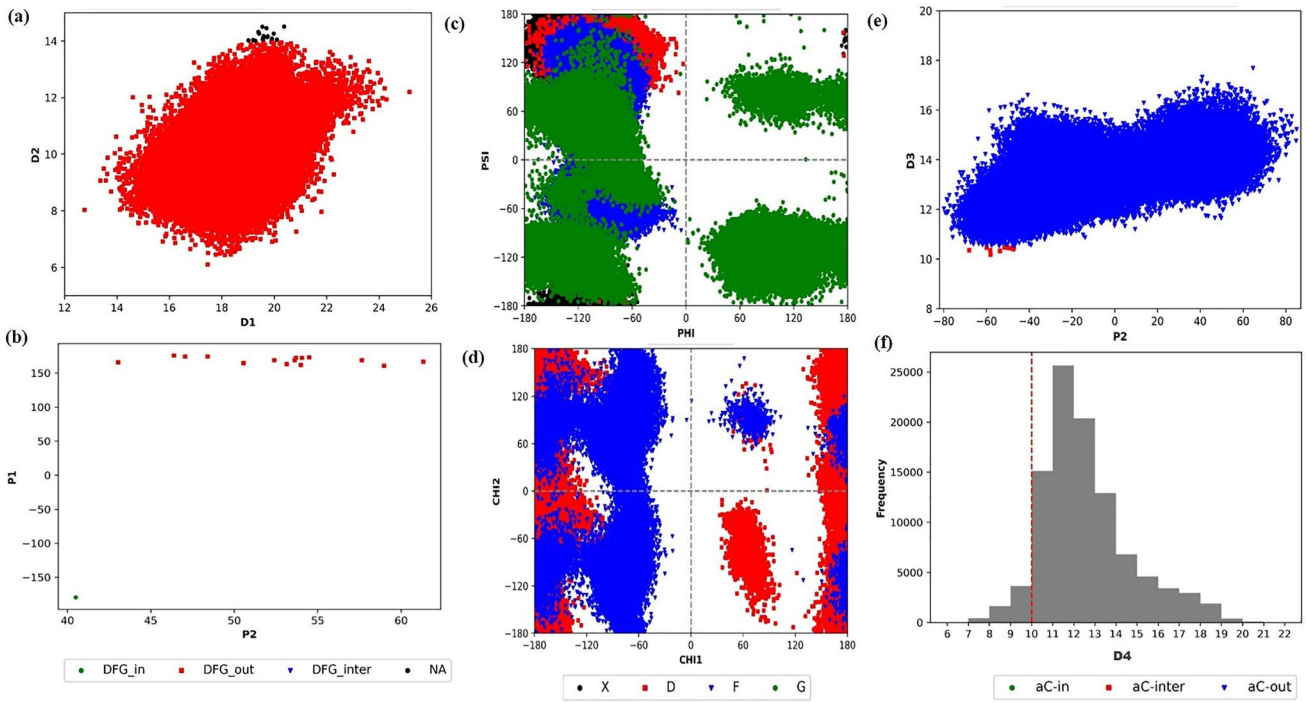

Fig. 3 Conformational classification of 5NXD- ($DFG_{out}\alpha C_{out}$) trajectory structures; DFG classification based on distance (**a**) D1 vs D2 (**b**) P2 vs P1; DFG dihedral classification (**c**) PHI vs PSI (**d**) CHI1 vs CHI2; αC classification of the DFG classified structures (**e**) P2 vs D3 (**f**) D4

"DFG" positioning as that of CIDI conformation, however, these were observed to have varying orientations. On the other hand, the other classified intermediates namely, CinterDI and CinterDinter, either of the salt-bridge forming residues (K360 and E376) were found to be outward-facing; unlike the CODO conformation wherein, both these residues are observed to be drastically facing away from each other, thereby incompatible for the salt-bridge formation. These classified conformations were observed to have variations in their respective catalytic and regulatory spine residues. In the case of the catalytic spine, all residues (A345, V358, L413, C457, L458, I459) were observed to have fluctuations, however, the long transitions were observed for the residues A345 (CinterDinter, CinterDI), V358 (CinterDinter), C457 (CIDI), I459 (CODI, CODinter) respectively, while L458 except for CODinter and CinterDI showed long transitions in other conformations (Table S4). In these residues, the high PAD degree for long transition was observed in CIDI conformation for the residue C457 (71.68°), followed by L458 in CIDinter (64.15°). The residues F391 and H449 of the regulatory spine (M380, K390, F391, H449, and F470) were observed to undergo long transitions in all conformations except for CIDI and CIDinter. Amongst the fluctuating residues of the regulatory spine, the highest PAD degree was observed in CinterDinter specifically for the residues K390 (71.99°) and F470 (151.32°). In the case of the XDFG motif, all residues were observed to have long transitions in CinterDinter conformation, whereas D469 was found to have fluctuation only in conformation (37.66°). However, D469 showed short transitions in CIDI (49.22°), CIDinter (88.66°), and long transitions in other conformations. The positioning of F470 residue that determines the active and inactive of kinase was observed to have high fluctuations in all these classified conformations, with a high PAD degree observed in CinterDinter (151.32°), while the short transition was observed only in CODinter (67.64°).

Overall, the classification of the respective MD trajectory resulted in the following states: $DFG_{in}\alpha C_{in}$ [CIDI] (Active state), and the intermediate/inactive states namely, $DFG_{out}\alpha C_{in}$ [CIDO], $DFG_{out}\alpha C_{inter}$ [CinterDO], $DFG_{in}\alpha C_{inter}$ [CinterDI], $DFG_{in}\alpha C_{out}$ [CODI]; $DFG_{inter}\alpha C_{in}$ [CIDinter], $DFG_{inter}\alpha C_{inter}$ [CinterDinter], $DFG_{inter}\alpha$Cout [CODinter] and the inactive state**:** $DFG_{out}\alpha C_{out}$ [CODO] which were considered for further analysis.

Identifying potential effectors by perturbation response scanning

The perturbation-response scanning (PRS) approach was employed to enumerate the allosteric effect of each residue on all other residues in the conformations in response to an external perturbation. First, the PRS map provides information on the sensitivity and effect of a given residue in transmitting signals. Then, two groups of residues including

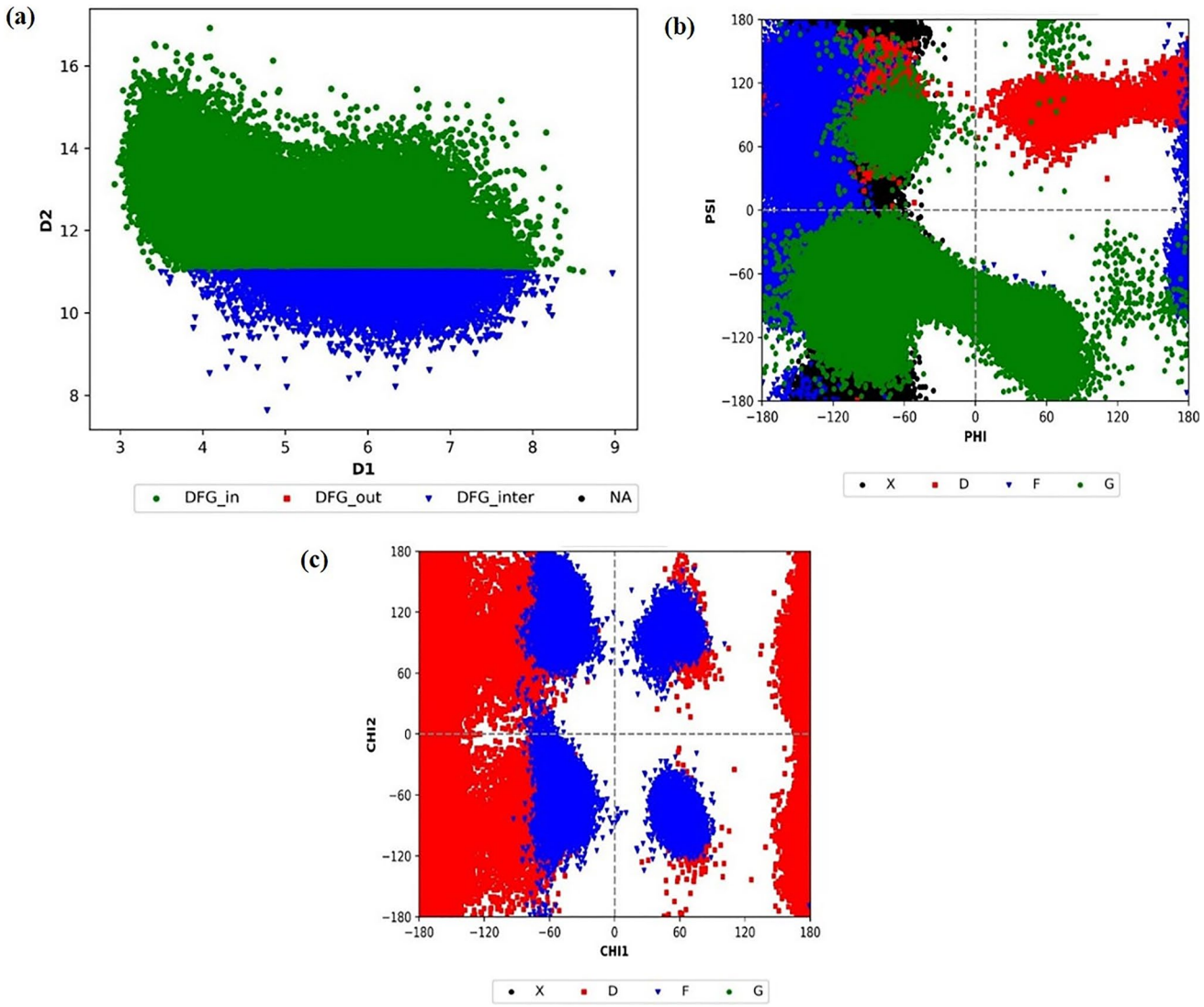

Fig. 4 DFG classification of LIMK2 ($DFG_{in}\alpha C_{in}$) based on distance (**a**) D1 vs D2; and dihedral (**b**) PHI vs PSI (**c**) CHI1 vs CHI2

sensors and effectors can be predicted from this map, which is potentially involved in allosteric signal sensing and transmission. In Figs. 6 and 7, the PRS map highlights some key residues that can be served as the strongest sensors and effectors in the kinase domain of LIMK2. Expect for the CODO conformation, the strongest effectors were observed to span at the N-lobe that constitutes the ATP binding cavity. In the case of CODO conformation (Table 2), the effectors were observed to be clustered at the C-lobe attributed to the substrate binding. Among these conformations, the CIDinter, CinterDI, CODinter, and CODI were found to have strong effectors than the intermediate conformations especially spanning the N-lobe of the kinase domain.

While the CIDI conformation showed the least effectors and sensors than the other conformations (Table 2). Based on the sensory profiles, it was inferred that the strong sensors were mainly the residues conferring the flexible loop regions that connect the β sheets spanning the conserved residues of the kinase domain. In the case of the effector profiles, mostly the β sheet conferring residues were observed to be the strongest effectors, amongst these, the β8 that precedes the XDFG motif was found to be a prominent effector in CinterDO, CIDinter, CinterDI, and CODO, whereas A468 ("X"-DFG) was observed as strong effector only in the intermediate conformation CODinter. Moreover, the crucial αC residue F373 (CODinter) and L374 (CIDinter) followed by αD residue L413 were shown to be strong effectors in CinterDI and CIDinter. Through PRS map observation, it was inferred that among the intermediates, the CIDinter conformation had the maximum number of conserved kinase domain residues as the effectors. Moreover, the overall PRS analysis reveals that the conserved β-sheet, G/P.loop,

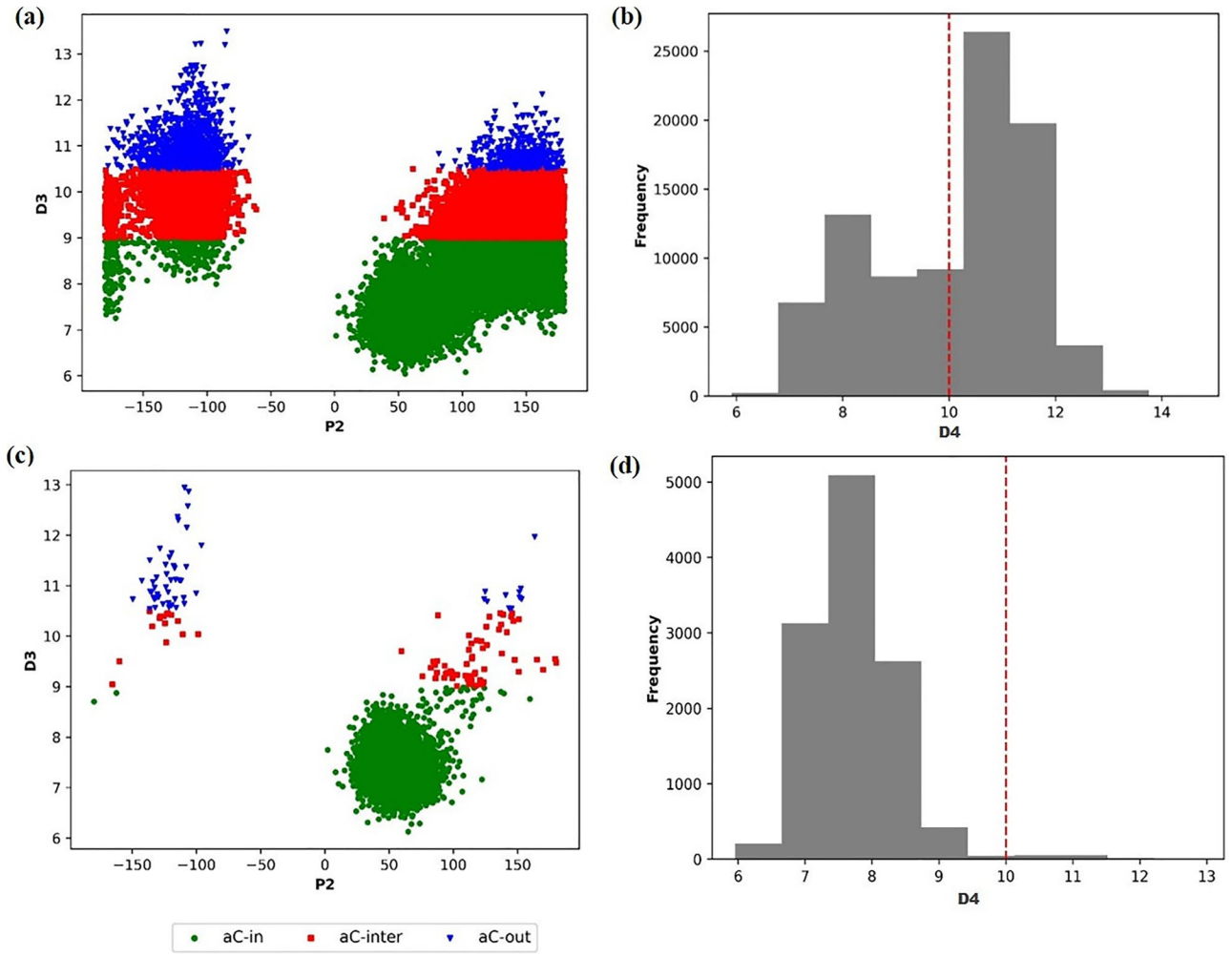

Fig. 5 αC classification of the DFG classified structures; DFG_{in} classified structures (**a**) P2 vs D3 (**b**) D4; DFG_{inter} classified structures (**c**) P2 vs D3 (**d**) D4

and α helices (αC and αD) to be as prominent effectors. Furthermore, the effectors, enriched in these intermediate conformations were presumably considered to be affecting its folding and stability, thereby favoring conformational transitions.

Structural network analysis of the intermediate states with the active and inactive states of classified structures

On observing the I_{min} which defines the lowest interaction strength between the links, it was observed that among the conformations sampled, the CIDI (active state) had the least I_{min} value of 3.83, wherein the CODO (4.27) and the intermediates showed I_{min} ranging > 4. However, on comparing the I_{min} value of the intermediates between CIDI and CODO, it was inferred that CIDI has drastic network changes with the lowest interaction strength, while the CODO features negligible differences. This defines that the interaction forces were lower in CIDI (active state) conformation than the intermediate and inactive conformations, which infers that there is a greater number of residue rearrangements, disrupting intra-molecular networks and interactions (Table S5). Therefore, these observations indicate that conformational variations affect the intramolecular interaction pathways and affect the structural stability of LIMK2 to a greater extent as per the PSN-ENM approach (Table S5).

Overall comparative global network and metapath analysis of the intermediates to CIDI and CODO conformations, among the intermediates, CODinter, CIDinter, and CinterDinter conformations have the least number of specific nodes, thereby indicating that these conformations share a higher proportion of shared nodes with the CIDI conformation. In the case of the CODO conformation, the CinterDO and CIDO were observed to have the least number of specific nodes. Hence, it reveals that the number of nodes and

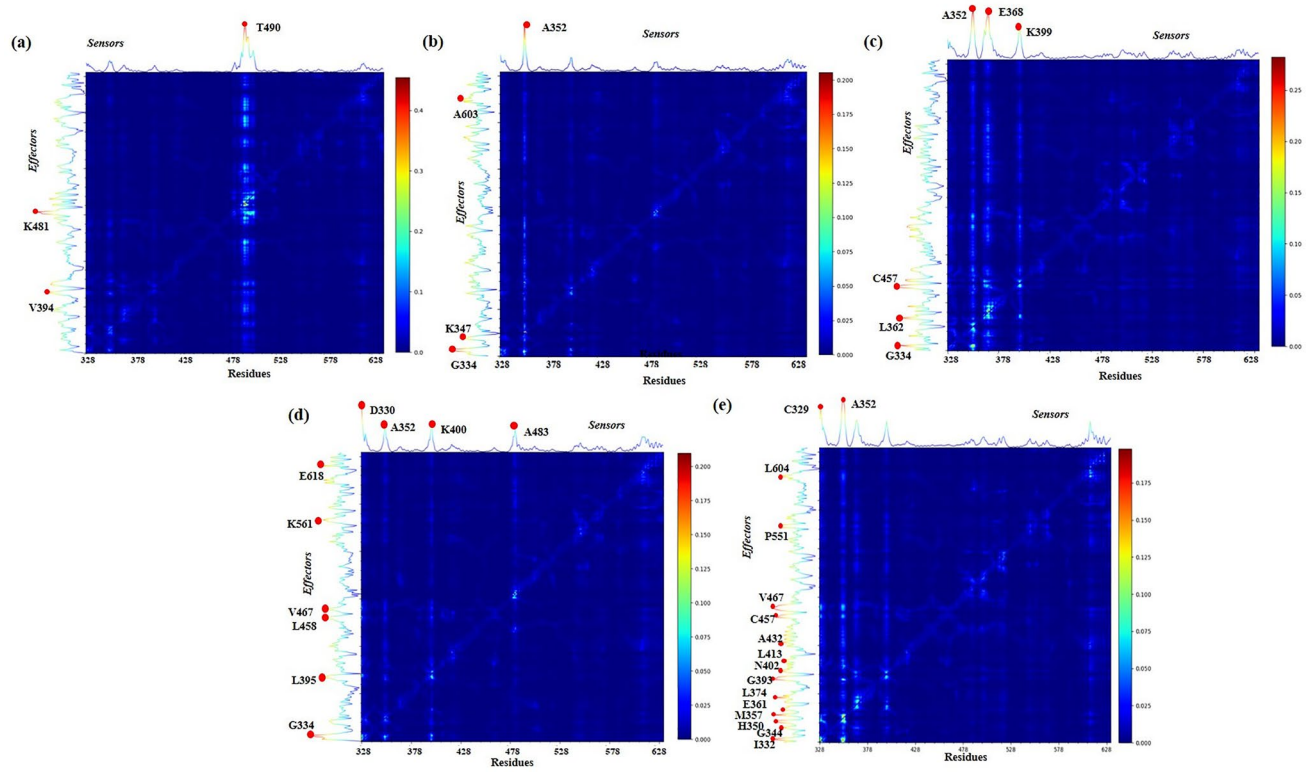

Fig. 6 The PRS map of the conformational states (**a**) CIDI (**b**) CIDO (**c**) CinterDinter (**d**) CinterDO and (**e**) CIDinter. The peaks correspond to the scale that points to efficient propagators of signals or effectors

links of intermediates that share a higher percentage of shared nodes with the CIDI and CODO conformations may correspond to intermediates that relate to the two states. On analyzing the total number of metapaths, CIDO followed by CODinter has a higher number of metapaths than CODO, while CinterDinter was observed to have higher metapaths compared to CIDI conformation. On comparing CIDI and CODO, the CIDI showed a higher number of metapaths (Tables S6 and S7).

Since αC and DFG define the conformational state of kinases, the metapath specific to the allosteric signalling that mediates along with αC and DFG in all the conformations was determined. In the case of the shared links of the CIDI state along with the intermediate states, it was observed that CinterDinter had a maximum of four links followed by CIDinter and CODinter sharing two and one link respectively, with CIDI. The intermediates CIDinter and CODinter share the following links with CIDI that connect L442 (αE) → I447 (β6), E376→ F470 and L442 → I447, respectively. The intermediate CinterDinter shares nearly four links with CIDI that connect the following nodes H449 (HRD) → L452 (C.l) → N456 (C.l) → D469 (DFG) and R482 (A.l) → E479 (A.l) (Table S8, Table 3 and Figs S6–S8). In terms of shared nodes, apart from αC-E376 and D469 (DFG), the intermediates CinterDinter, CIDinter, and CODinter shared seven, four, and three nodes respectively with CIDI, while both CinterDO and CIDO shared a single node with CIDI. On observing the shared nodes, it was inferred that H449 of the HRD motif were shared nodes in the intermediates CinterDinter, CIDinter, and CinterDinter, while F470 of the DFG motif was a shared node in CIDinter, CODinter, respectively. Apart from these nodes, the intermediate CinterDinter conformation that shares maximum nodes with CIDI comprises Y441 (αE), L452, and N456 of C.l, E479, and R482 of A.l. While both the CIDinter and CODinter conformation were observed to have the same shared nodes I447 (β6), and L442 (αE) with CIDI conformation. Whereas, CinterDO and CIDO shared the nodes L442 (αE), and R450 (HRD) with CIDI, respectively.

On comparative analysis of the prominent nodes of intermediates along with CIDI conformation that features the conserved structural elements of the kinase domain, only CIDI conformation was observed to have higher conserved residues as prominent nodes than the intermediate conformations. In CIDI, the highly conserved motifs HRD (H449, R450) and DFG residues (F470, G471) and the connecting loop residues namely, C.l (L452, N456), A.l (L472, S473), and b.l (V388) were found to be the prominent nodes for allosteric communication in the filtered metapath. Apart from these residues, the αE residues Y441, L442, H443,

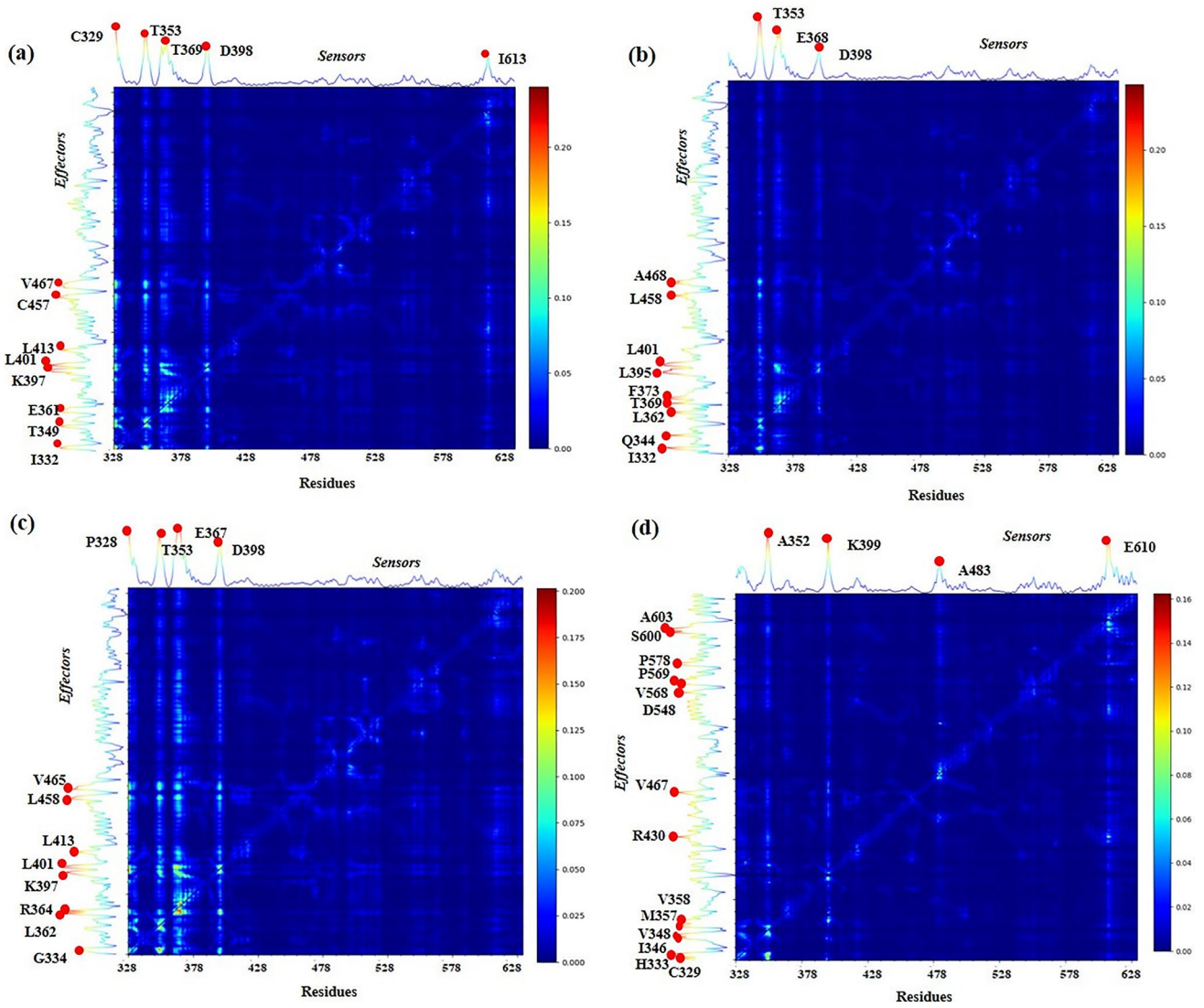

Fig. 7 The PRS map of the conformational states (**a**) CinterDI (**b**) CODinter (**c**) CODI and (**d**) CODO. The peaks correspond to the scale that points to efficient propagators of signals or effectors

S444, and β6 residues C446, I447, and I448 were found to have higher dominance in the CIDI metapath. In the case of intermediates, DFG residues, F470 (CinterDinter), G471 (CinterDinter, CIDinter), and the K360 of β3 which facilitates salt bridge formation were found to be significant nodes in the intermediates CODinter and CinterDI. Among the conserved α helices, the αC residues, F373 (CODinter, CinterDI), M380 (CinterDinter); αE residues H443 (CIDO), M445 (CinterDinter), and L413 of αD (CODI) were observed to be involved in the allosteric communication in their respective intermediates. Of the intermediates, the CODI conformation showed a higher predominance of C.l residues S454, N453, and N456 as nodes, while the H449 of the HRD motif acts as a node in the metapath only in intermediates CinterDO, and CIDO. Whereas the β–sheets spanning residues were highly profound as nodes in intermediates CODinter, CODI and CinterDO followed by CinterDI.

From the total number of metapaths identified in the intermediates along the CIDI, the CinterDinter had a maximum of four metapaths while CIDinter and CIDI showed two metapaths. In CIDI the metapaths connected αC to DFG were mediated either via I448 (β6)/ L442 (αE) or T526/ I476 (A.l), while in the case of CIDinter the highly conserved residues namely, H449 (HRD)/I447 (β6) facilitated the metapaths connectivity. However, in the intermediate CinterDinter, the nodes L495/R482 leading to varying connectivity along the metapaths are specified as follows: L495→ K492 → H449 (HRD)/ N456 (C.l) and R482→ M485 → H449 (HRD)/ N456 (C.l). This observation suggests that the HRD motif and C.l residues could act as crucial nodes for the

Table 2 PRS analysis of the conformational states; **The loop conferring residues of conserved motifs are coloured in red, while the residues of α helices are shown in green and the "X" –DFG are in blue

Conformations	Effectors	Sensors
CIDI	**V394**,K481	T490
CIDO	**G334,K347**,A603	A352
CinterDinter	**G334, L362, C457**	A352,E368,K399
CinterDO	**G334,**L395,**L458,V467**,K561,E618	D330, A352,K400,A483
CIDinter	I332,**G344**,H350,**M357, E361,L374,G393,N402,L413**,A432, **C457,V467**,P551,L604	C329,A352
CinterDI	I332,T349,**E361**,K397,L401,**L413,C457,V467**	C329,T353,T369,D398, I613
CODinter	I332,**Q344,L362**,T369,**F373**,L395,L401,**L458,A468**	T353, E368,D398
CODI	**G334,L362,**R364,K397,**L401,L413,L458,V465**	P328,T353,E367,D398
CODO	C329,H333,**I346,**V348,**M357,V358**,R430,**V467,**D548,V568, P569,P578,S600, A603	A352,K399,A483,E610

Table 3 Filtered metapath connecting αC and DFG of intermediates with CIDI (active state)

Conformational states	Metapath connecting αC to DFG
CIDI	E376 –> G471 –> I448/L442 –> R450 –> S473 –> D469
CIDO	E376 –> R496 –> R450 –> R493 –> E525 –> H443 –> H449 –> R474 –> D469
CIDI	E376 –> F470 –> I447/L442 –> H449 –> L472 –> D469
CinterDI	E376 –> F373 –> L403 –> F391–> K360 –> D469
CIDI	E376 –> G471 –> I448/L442 –> R450 –> Y523 –> D469
CinterDO	E376 –> I476 –> I447–> L442 –> H449–> R474 –> D469
CIDI	E376 –> F470 –> I447/L442 –> H449 –> L472 –> D469
CODI	E376 –> F373–> L403 –> F391 –> L389 –> K460 –> V466 –> N387 –> T464 –> F431 –> V465 –> L413 –> S454 –> Q452 –> E538 –> N453 –> N456 –> D469
CIDI	E376 –> F470 –> I447–> L442 –> V388–> H386 –> Y441–> R496 –> E479–> R482–> T526/I476–> I448 –> H443–> H449 –> L452–> N456 –> D469
CinterDinter	E376 –> G471 –> M380–> F470–> L383 –> Y441–> M446 –> E479–> L495/R482
	R482–> M485–> E525–> V527–> F594–> M439–> L452–> H449/N456–> D469
	L495–> K492–> E525–> V527–> F594–> M439–> L452–> H449/N456–> D469
CIDI	E376–> F470–> I447–> L442–> V388–> H385–> Y441–> R396–> S444–> E479–> M485–> T526–> I448–> H443 –> H449–> L452–> N456–> D469
CODinter	E376–> L472–> I447–> L442–> F470–> M380–> F391–> L403–> F373–> L362–> G343–> K360–> D469
CIDI	E376–> F470–> I447–> L442–> H385–> Y441–> R496–> S444–> E479–> R482–> T526–> I448–> H443–> H449 –> L452–> N456–> D469
CIDinter	E376–> F470–> V379–> L383–> L442–> H449/I447–S473–> D469

allosteric communications along with the intermediates in the metapaths, in specific, H449 was found to be the prominent node in both intermediates and CIDI conformation.

It was observed that the CODO (inactive state) in comparison with other intermediate conformations does not have any shared links with most of the intermediates, except for the intermediates CIDinter and CinterDO that have shared links with CODO namely, R474 (A.l) → D469 ("D"-DFG). However, in terms of shared nodes, except for αC-E376 (αCE) and D469 of DFG, the intermediates, CODinter shares two nodes namely, F373 (αE) and S473 (A.l). While the other intermediates CODI and CinterDO share a single node spanning the A.l i.e., S473 and R474, respectively with CODO. Besides this, on observing the well-known conserved structural features of the kinase domain that are involved in allosteric signalling, it was inferred that in CODO conformation, only F373 (αC) and S473 (A.l) were the only prominent nodes that favored the allosteric communication of the αC and DFG. Among the intermediates, the highly conserved HRD motif of the catalytic loop (C.l) residues namely, H449 and R450 were found to be the prominent nodes of the metapath of CIDO, CIDinter, CinterDO,

and CinterDinter conformations. Apart from the HRD motif, the CinterDinter conformation was found to have other residues of the C.l Viz, D451, N453, and N456 that aid in allosteric communication, while N456 was observed as a crucial node in both CODI and CODinter among the intermediates. Of the conserved α helices of the kinase domain, in the case of αE, the residues H443 (CIDO, CIDinter), Y441 (CODinter), L442 (CODI, CinterDO, CODinter); wherein the αC residues S382 (CODinter, CODI), M380 (CODI, CIDinter); αD residues D415 (CODinter, CODI) and R418 (CinterDO) were observed to be the significant nodes for the allosteric communication. Other than these elements, only CODI conformation was observed to have hinge residues namely, Y407, E406, and F470 (F-DFG), while the Linker residue T412 was found to be in the CODinter and CODI metapath. Overall, from the filtered metapath, among the intermediates, CODinter and CODI intermediate conformations were observed to have defined the metapath connecting a higher degree of conserved structural elements of the kinase domain. In specific, most of the residues spanning the conserved β sheets and the loops namely b.l, G/P.loop were observed to be involved in the metapath for allosteric communication. The other intermediates namely, CinterDO and CinterDinter have β sheet residues as prominent nodes in the metapath. (Table S8, Table 4 and Figs S9–S11).

On observing the filtered metapaths, the CODO, CODinter, CODI, and CinterDinter conformations have shown a maximum number of metapaths i.e. 2. Whereas, the major shift in the path of CODO was observed either via R364 or L475 that connects to D469 and when compared to the intermediates, CODO has the least number of nodes that connects to D469. Amongst the intermediates, the altered nodes connecting the metapath to D469 are as follows, CODinter: N456 (C.l)/M359 (β3), CODI: V358 (β3)/N456 (C.l), CinterDinter: D451/N456 of C.l. This infers that the C.L and β3 as the prominent nodes that intermediates the allosteric signalling, in the specific residue (N456) aids as the prominent node. In comparison, it was observed that the inactive CODO state has the shortest metapath than the other intermediates.

Inactive state conformation classification

On analyzing the A.l classification of the sampled inactive conformations of LIMK2 (Fig. S12–S13, Table 5), CIDO conformations failed to surpass the descriptor one criteria when compared to other conformations. Hence, were considered for further A.l classification. Moreover, it was observed that except for CODI and CODinter, all the other intermediate inactive conformations were found to have Open DFG-out type (OD) of A.l. The A.l is, similar to DFG-in structures, in an unfolded conformation in the "open DFG-out" conformation which opens up the active site and is compatible with the binding of Type1 inhibitors. Collapsed G/P.l conformations might prefer an A.l in the open DFGout-state compared with closed type 2 and A-under-P conformations. While the CODO conformation was found to have a closed type 2 state of A.l, which infers the inactive state of the kinase with A.l is completely intact and closed and hence making it non-compatible for substrate binding.

While classifying the intermediate inactive conformations based on the G/P. loop (Figs. S14–16, Table 6), CIDO conformation was the only one found to surpass all four criteria and was found to have collapsed G/P.

Table 4 Filtered metapath connecting αC and DFG of intermediates with CODO (Inactive state)

Conformational states	Metapath connecting αC to DFG
CODO	E376 –> L475/R364–> R496 –> S473–> R474 –> D469
CinterDO	E376 –> I476 –> I447 –> L442–> H449 –> R474–> D469
CODO	E376 –> L475/R364–> R496 –> S473–> R474 –> D469
CIDinter	E376 –> M380–> Y441–> R552–> E486–> T489 –> H449–> H443 –> R450–> R474–> D469
CODO	E376 –> R364/L475 –> R496 –> S473 –> R474–> D469
CinterDI	E376 –> R418–> L404–> F391 –> V358 –> D469
CODO	E376 –> F373–> R364 –> S473 –> D469
CIDO	E376 –> R496–> R450 –> R493 –> E525–> H443 –> H449 –> R474 –> D469
CODO	E376 –> F373–> R364 –> S473 –> D469
CinterDinter	E376 –> L472–> L475–> H449–> I448–> V477–> M439–> N453–> D451/N456–> D469
CODO	E376 –> R364/L475 –> R496–> S473–> R474–> D469
CODI	E376–> S473–> I447–> C446–> S382–> Y441–> V388–> L442–> L383–> M380–> E406-> Y407–> V358–> Q344–> G338–> A345–> M357–> F470–> D415–> T412–> L332–> V358/N456–> D469
CODO	E376 –> F373–> R364–> S473–> D469
CODinter	E376 –> L475–> I447–> C446–> L383–> Y441–> V388–> L442–> S473–> S382 –> F391–> L403–> F373-> E361–> Q344–> G334–> I346–> V336–> D415–> T412–> G338–> N456/M359–> D469

Table 5 Activation loop (A.l) classification for the intermediate and inactive states of LIMK2; ** (green tick) if the conformations satisfy the conditions, (red cross) if the conformations does not satisfy the conditions

Intermediate/ Inactive states	$xi_{DFG\{-1,D\}}$ [P1 = ξ_0 (A468,D469,F470,G471)]	$xi_{DFG\{F,G\}}$ [P2 = ξ_0 (D469,F470,G471,L472)] and Cα-Cα [D5 = Cα M445- Cα S474]		
		Close Type 2 (CT2)	Open DFG-out (OD)	Closed A-under-P (CAP)
CIDO	✘	-	-	-
CinterDO	✓	✘	✓	✘
CODO	✓	✓	✘	✘
CinterDI	✓	✘	✓	✘
CODI	✓	✘	✘	✘
CIDinter	✓	✘	✓	✘
CinterDinter	✓	✘	✓	✘
CODinter	✓	✘	✘	✘

loop conformation. While other conformations were found to have surpassed only three criteria. Among these, CinterDO was observed to have collapsed the state of G/P. loop, wherein the CODinter showed the Stretched state of the G/P.loop. Apart from these conformations, all the other conformations were observed to have both collapsed and stretched states of the G/P.loop. The variations in G/P. loop are essential in accommodating the substrate. Based on previous studies, it was inferred that the collapsed G/P. loop has a higher predominance for inhibitor or substrate binding [45], thereby it indicates that all the classified conformations are highly suitable for substrate binding.

Druggable site analysis

On analyzing the druggable pockets/cryptic sites of the classified LIMK2 active and inactive states (Tables 7, 8 9). In the active CIDI of LIMK2, Pocket2 with a higher druggability score and high volume was inferred as the most feasible and novel druggable site. On comparative analysis with the KLIFS orthosteric and allosteric ligands (Tables S10–S23, Figs. 8 and 9, Fig. S17), Pocket2 (CIDI) was observed to accommodate a vast range of Type 1, 1.5 inhibitors that are in clinical trials. The other kinase inhibitors in clinical trials such as Type 2 inhibitors namely LIG ID: 0LI and STI and Type3 inhibitors namely, LIG ID: EUI was found to be overlaid at the Pocket2 of CIDI conformation. This pocket covers the overall N-lobe ATP binding cavity inclusive of the XDFG motif, thereby it aids in the accommodation of

Table 6 G/P.loop classification inference for the Intermediate and inactive states of LIMK2; ** Green tick Collapsed [CL], – not satisfies CL condition, red cross Stretched [ST], – not satisfies ST condition

Inactive conformations	Cα-Cα [D6 =Cα(342F)-Cα(455H)]		$psi_{G\text{-motif-1}}$ [T1 = (Φ, Ψ) 337]		$psi_{G\text{-motif+1}}$ [T2 = (Φ, Ψ) 344]		$xi_{G\text{-motif}\{+1,+2\}}$ [P3= ξ_0 (343,344,345,346)]		Inference **satisfy three among the four conditions
P-Loop Types	CL	ST	CL	ST	CL	ST	CL	ST	
CIDO	✓	-	✓	✘	✓	✘	✓	✘	**Collapsed**
CinterDO	✓	-	✓	-	-	✘	✓	✘	**Collapsed**
CODO	✓	-	✓	✘	✓	✘	✓	✘	**Collapsed, stretched**
CinterDI	✓	-	✓	✘	✓	✘	-	✘	**Collapsed, stretched**
CODI	✓	-	✓	✘	✓	✘	-	✘	**Collapsed, stretched**
CIDinter	✓	-	✓	✘	✓	✘	-	✘	**Collapsed, stretched**
CinterDinter	✓	-	✓	✘	✓	✘	-	✘	**Collapsed, stretched**
CODinter	✓	-	-	✘	-	✘	-	✘	**stretched**

Table 7 Druggable pockets identified from the classified conformations

Conformations	CIDI		CinterDO-OD		CinterDI-OD		CIDinter-OD			CinterDinter-OD		CODO-CT	
Parameters	Pocket1	Pocket2	Pocket1	Pocket3	Pocket1	Pocket2	Pocket1	Pocket4	Pocket6	Pocket2	Pocket3	Pocket1	Pocket9
Druggability Score	0.59	0.74	0.67	0.50	0.86	0.55	0.68	0.68	0.54	0.63	0.52	0.92	0.51
Number of Alpha Spheres	59.00	185.00	60.00	40.00	42.00	21.00	50.00	40.00	34.00	42.00	32.00	77.00	43.00
Total SASA (Å^2)	64.79	488.71	70.92	108.13	123.20	117.58	184.98	124.40	100.76	161.85	95.32	92.99	120.68
Polar SASA (Å^2)	4.33	244.77	33.48	38.01	32.63	37.88	53.84	49.53	44.00	74.90	31.32	40.98	31.32
Apolar SASA (Å^2)	60.46	243.94	37.44	70.12	90.57	79.70	131.14	74.87	56.76	86.95	64.00	52.01	89.36
Volume (Å^3)	309.38	1752.60	314.87	374.32	619.33	406.88	607.69	443.21	229.51	507.53	277.78	408.81	332.80
Mean local hydrophobic density	56.00	27.19	17.58	16.89	21.55	7.69	19.33	20.36	16.84	11.60	18.00	32.27	31.00
Mean alpha sphere radius (Å)	3.89	3.95	3.66	3.94	4.12	3.86	3.78	3.91	3.62	3.97	3.82	3.72	3.86
Mean alpha. sphere. solvent access (Å)	0.51	0.48	0.42	0.53	0.71	0.64	0.52	0.48	0.34	0.56	0.49	0.39	0.44
Apolar alpha sphere proportion	0.97	0.35	0.32	0.45	0.74	0.62	0.48	0.55	0.56	0.36	0.59	0.48	0.74
Hydrophobicity score	62.93	21.29	16.47	34.80	33.00	18.00	26.24	17.08	22.79	-1.85	35.67	48.06	45.75
Volume score	4.21	3.74	4.00	4.50	4.78	4.56	4.65	4.39	4.21	3.85	3.78	5.00	4.58
Polarity score	0.00	19.00	10.00	5.00	5.00	6.00	9.00	8.00	8.00	9.00	4.00	7.00	5.00
Charge score	0.00	-2.00	2.00	2.00	3.00	2.00	-2.00	1.00	1.00	1.00	1.00	1.00	3.00
Proportion of polar atoms	9.38	42.06	38.46	39.29	30.30	34.78	37.50	34.48	39.13	40.00	38.10	32.43	28.57
Alpha sphere density	3.05	9.43	3.95	3.83	6.14	5.73	6.47	4.63	3.73	4.44	3.15	5.06	3.45
Cent. of mass—Alpha Sphere max distance (Å)	6.61	24.55	8.00	10.71	14.46	14.71	12.95	11.27	8.96	11.70	6.52	12.29	9.18

Table 8 Druggable pockets identified from the classified conformations

Conformations	CIDO-CL	CinterDI-CL		CIDinter-CL		Cinter Dinter-CL	CODO-CL	CIDinter-ST				CODI-ST	CODO-ST
Parameters	Pocket1	Pocket3	Pocket6	Pocket1	Pocket9	Pocket9	Pocket7	Pocket1	Pocket2	Pocket11	Pocket15	Pocket 1	Pocket 1
Druggability score	0.52	0.71	0.52	0.81	0.65	0.56	0.60	0.85	0.59	0.52	0.60	0.68	0.79
Number of alpha spheres	42.00	19.00	16.00	65.00	36.00	25.00	137.00	43.00	35.00	37.00	28.00	47.00	78.00
Total SASA (Å^2)	55.51	105.92	96.13	110.93	234.05	128.82	438.56	194.51	97.46	196.47	175.26	140.46	114.89
Polar SASA (Å^2)	20.49	32.25	52.66	21.56	153.14	57.41	215.15	52.75	33.46	91.41	81.07	37.66	40.02
Apolar SASA (Å^2)	35.02	73.67	43.48	89.36	80.91	71.41	223.41	141.77	64.00	105.06	94.20	102.81	74.87
Volume (Å^3)	247.57	389.27	418.79	448.57	702.50	402.68	1517.63	644.78	297.29	586.16	483.60	506.43	417.69
Mean local hydrophobic density	23.00	12.00	5.00	34.82	6.73	15.00	24.20	17.41	18.00	11.00	6.00	16.08	26.54
Mean alpha sphere radius (Å)	3.79	3.91	4.14	3.82	3.90	4.02	3.86	3.92	3.93	3.99	3.94	3.85	3.84
Mean alpha. sphere. solvent access (Å)	0.46	0.52	0.74	0.47	0.55	0.59	0.47	0.57	0.42	0.57	0.54	0.52	0.49
Apolar alpha sphere proportion	0.57	0.68	0.38	0.60	0.31	0.64	0.50	0.63	0.54	0.38	0.36	0.55	0.47
Hydrophobicity score	52.44	25.56	17.00	19.74	14.38	37.91	42.84	5.67	14.36	32.87	-1.46	29.71	20.60
Volume score	4.75	4.78	4.83	4.00	4.31	4.36	4.62	4.80	4.07	4.60	4.09	4.64	4.00
Polarity score	5.00	4.00	4.00	9.00	11.00	3.00	15.00	10.00	8.00	8.00	8.00	6.00	8.00
Charge score	2.00	0.00	1.00	2.00	-2.00	3.00	4.00	0.00	1.00	1.00	1.00	-3.00	2.00
Proportion of polar atoms	29.63	30.00	50.00	34.88	44.44	32.00	35.00	34.21	40.00	41.94	44.44	34.29	36.36
Alpha sphere density	3.11	4.88	4.03	4.62	7.64	4.45	9.32	6.76	3.46	5.34	5.25	5.53	5.07
Cent. of mass—Alpha Sphere max distance (Å)	7.02	10.47	8.99	9.99	20.03	11.42	22.44	14.56	7.03	12.80	14.01	11.83	11.68

Table 9 Residues spanning the druggable Pocketome of the classified conformations

Conformations	Pockets	Pocket residues
CIDI	Pocket1	429VAL,432ALA,536LEU,540ILE,560VAL,563PHE,573PRO,574PRO,576PHE,579LEU,580ALA,583CYS,601PHE,604LEU
	Pocket2	337LEU,338GLY,339LYS,340GLY,343GLY,344GLN,345ALA,358VAL,359MET,360LYS,361GLU,362LEU,368GLU,369THR,372THR, 373PHE,376GLU,380MET,388VAL,389LEU,405THR,406GLU,408ILE,410GLY,411GLY,412THR,414LYS,415ASP,418ARG,451ASP, 455HIS,456ASN,458LEU,466VAL,467VAL,468ALA,469ASP,470PHE,472LEU,508GLY,509ASN,510PRO
CinterDO-OD	Pocket1	440ALA,443HIS,444SER,479GLU,482ARG,493ARG,494THR,525GLU,526THR,527VAL,529ILE,583CYS,591ARG,592PRO,593ALA,594PHE,595SER
	Pocket3	553THR,554LEU,559ASN,560VAL,561LYS,581ALA,584CYS,585ARG,628MET,630TYR
CinterDI-OD	Pocket1	371LYS,374LEU,375THR,377VAL,378LYS,381ARG,382SER,391PHE,394VAL
	Pocket2	333HIS,334GLY,346ILE,361GLU,395LEU,396TYR,397LYS,400LYS,402ASN
CIDinter-OD	Pocket1	512TRP,513MET,514ALA,517MET,523TYR,524ASP,527VAL,528ASP,529ILE,555ASP,556PHE,585ARG,586LEU,587GLU,588PRO,590SER,591ARG
	Pocket4	490THR,491LYS,582ILE,585ARG,592PRO,596LYS,597LEU,600SER,618GLU,621GLU,622LEU,625THR,626VAL
	Pocket6	444SER,479GLU,480ARG,482ARG,483ALA,485MET,492LYS,494THR,495LEU,525GLU,527VAL,593ALA,594PHE,595SER
CinterDinter-OD	Pocket2	340GLY,414LYS,418ARG,453ASN,454SER,455HIS,538GLU,542GLN,543VAL,544TYR,545ALA,546ASP,547PRO
	Pocket3	553THR,558LEU,559ASN,560VAL,561LYS,581ALA,626VAL,627SER,628MET
CODO-OD	Pocket1	512TRP,513MET,514ALA,517MET,523TYR,524ASP,527VAL,528ASP,529ILE,555ASP,556PHE,585ARG,586LEU,587GLU,588PRO,590SER,591ARG
	Pocket9	489THR,490THR,491LYS,587GLU,589GLU,590SER
CIDO-CL	Pocket1	360LYS,360LYS,377VAL,377VAL,380MET,381ARG,383LEU,383LEU,383LEU,388VAL,389LEU,391PHE,391PHE,405THR,442LEU,447ILE, 467VAL,468ALA,468ALA,469ASP,469ASP,474ARG,474ARG,474ARG,476ILE,476ILE,476ILE
CinterDI-CL	Pocket3	416PHE,420MET,421ASP,423PHE,424PRO,426GLN,427GLN,431PHE,463LYS
	Pocket6	378LYS,381ARG,382SER,384ASP,391PHE,392ILE
CIDinter-CL	Pocket1	440ALA,443HIS,444SER,479GLU,482ARG,483ALA,485MET,492LYS,494THR,495LEU,525GLU,527VAL,529ILE,588PRO,591ARG,592PRO,593ALA, 594PHE,595SER
	Pocket9	433LYS,490THR,493ARG,579LEU,582ILE,595SER,596LYS,598GLU,599ASP,600SER,602GLU,618GLU,619LEU,621GLU,622LEU,626VAL
CinterDinter-CL	Pocket9	328PRO,329CYS,331LEU,350HIS,355LYS,357MET,381ARG,392ILE,393GLY,394VAL,404LEU
CODO-CL	Pocket7	337LEU,339LYS,340GLY,341PHE,342PHE,343GLY,344GLN,345ALA,358VAL,360LYS,361GLU,362LEU,363ILE,373PHE,376GLU,377VAL,380MET, 389LEU,400LYS,403LEU,405THR,407TYR,408ILE,410GLY,412THR,455HIS,458LEU,467VAL,468ALA,469ASP,470PHE,472LEU,474ARG,476ILE, 496ARG,503ARG,504TYR
CIDINTER-ST	Pocket1	485MET,491LYS,492LYS,513MET,517MET,524ASP,525GLU,555ASP,585ARG,586LEU,587GLU,588PRO,590SER,591ARG,630TYR
	Pocket2	440ALA,444SER,479GLU,491LYS,493ARG,494THR,525GLU,527VAL,529ILE,591ARG,592PRO,593ALA,594PHE,595SER
	Pocket11	375THR,378LYS,379VAL,381ARG,382SER,383LEU,384ASP,441TYR,445MET,446CYS,475LEU,476ILE,477VAL,478GLU,497LYS
	Pocket15	493ARG,494THR,582ILE,589GLU,590SER,591ARG,592PRO,593ALA,595SER,596LYS,618GLU
CODI-ST	Pocket1	513MET,514ALA,516GLU,517MET,523TYR,528ASP,529ILE,530PHE,555ASP,556PHE,586LEU,587GLU,588PRO,591ARG
CODO-ST	Pocket1	337LEU,338GLY,339LYS,342PHE,407TYR,412THR,451ASP,453ASN,455HIS,456ASN,470PHE,471GLY,474ARG,507VAL,508GLY

Type1, 1.5, 2, and 3 inhibitors. Pocket1 which spans the C-lobe constituting the α helices (αD and αE) showed an optimal druggable score and shall be considered as an allosteric site for further inhibitor identification studies. Both the

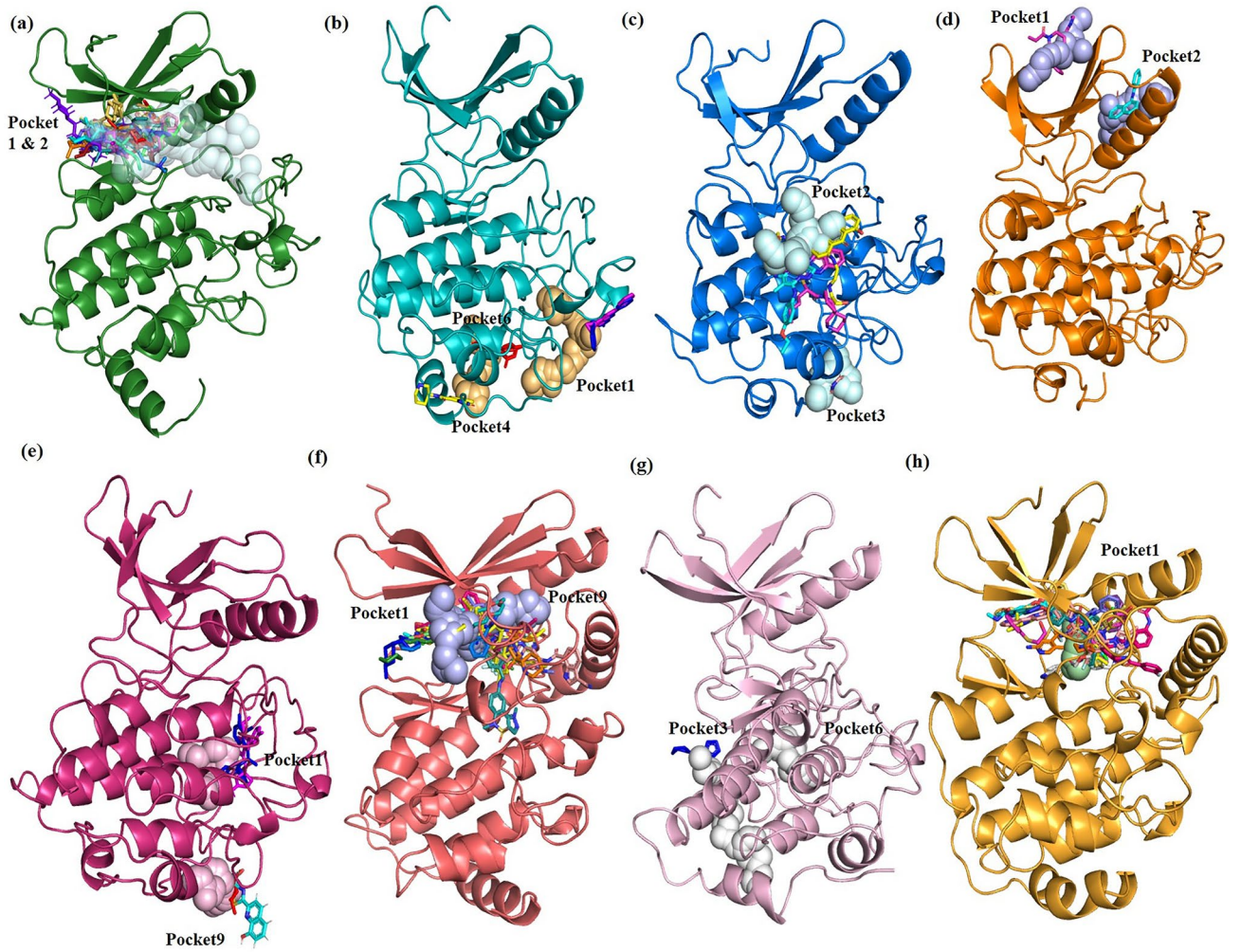

Fig. 8 Druggable pockets of the identified conformations and the overlaying respective orthosteric or allosteric ligands from KLIFS database, (**a**) CIDI (**b**) CIDinter-OD, (**c**) CinterDinter-OD (**d**) CinterDI-OD (**e**) CinterDO-OD (**f**) CODO-CT (**g**) CIDinter-CL and (**h**) CIDO-CL

pocket residues were observed to mediate the allosteric network in CIDI conformation.

In the case of intermediate inactive conformations classified based on A.l as OD state, the following intermediates CinterDinter, CIDinter, CinterDI, and CinterDO were predicted to have significant druggable sites as shown in Table 7. Amongst these, the CinterDO–OD conformation pockets both Pocket1 and Pocket3 are allosteric druggable pockets away from the ATP binding cavity and were found to overlap with the other allosteric ligand binding as follows, where Pocket3 overlaps with ligands (LIG ID)WG8 (PDB ID: 4D1Z) and HGQ (PDB ID: 6Q49), while Pocket1 overlaps with orthosteric ligands namely, G8H (PDB ID:6HK6), Q7K (PDB ID: 5M6U), Y5Y (PDB ID: 7LM2), respectively. The druggable pockets of CinterDI–OD, Pocket1 and Pocket2 are also observed to be distal from the ATP binding cavity. Moreover, Pocket1 resides at the cavity constituting the αC and β4 of the N-lobe, while Pocket2 covers the cavity formed by the conserved anti-parallel β-sheet of the N-lobe. Both these pockets were observed to be overlaid with many allosteric inhibitors, in specific, Pocket1 overlays with 2AN, a Type 3 inhibitor, whereas Pocket2 was found to be feasible in accommodating Type1.5 front inhibitor 1E8. All the predicted druggable pockets of CIDinter-OD were found to span at the C-lobe, amongst these only Pocket1 and Pocket4 overlap with other available allosteric inhibitors like 6RG(PDB ID: 4DLJ), IRG (PDB ID: 4EH2), and YDJ (PDB ID: 4W9W), W47 (PDB ID: 7JY4), respectively. In CinterDinter-OD both the allosteric pockets viz., the Pocket2 covers hinge, linker, and αD residues, while Pocket3 spans at the C-lobe were found to accommodate Type3 inhibitor HRZ (PDB ID: 6QA4).

The CODO-CT conformation was found to have overall higher druggability among all the conformational druggable sites and this pocket was observed to occupy the ATP binding cavity along the linker region. Moreover, these pockets

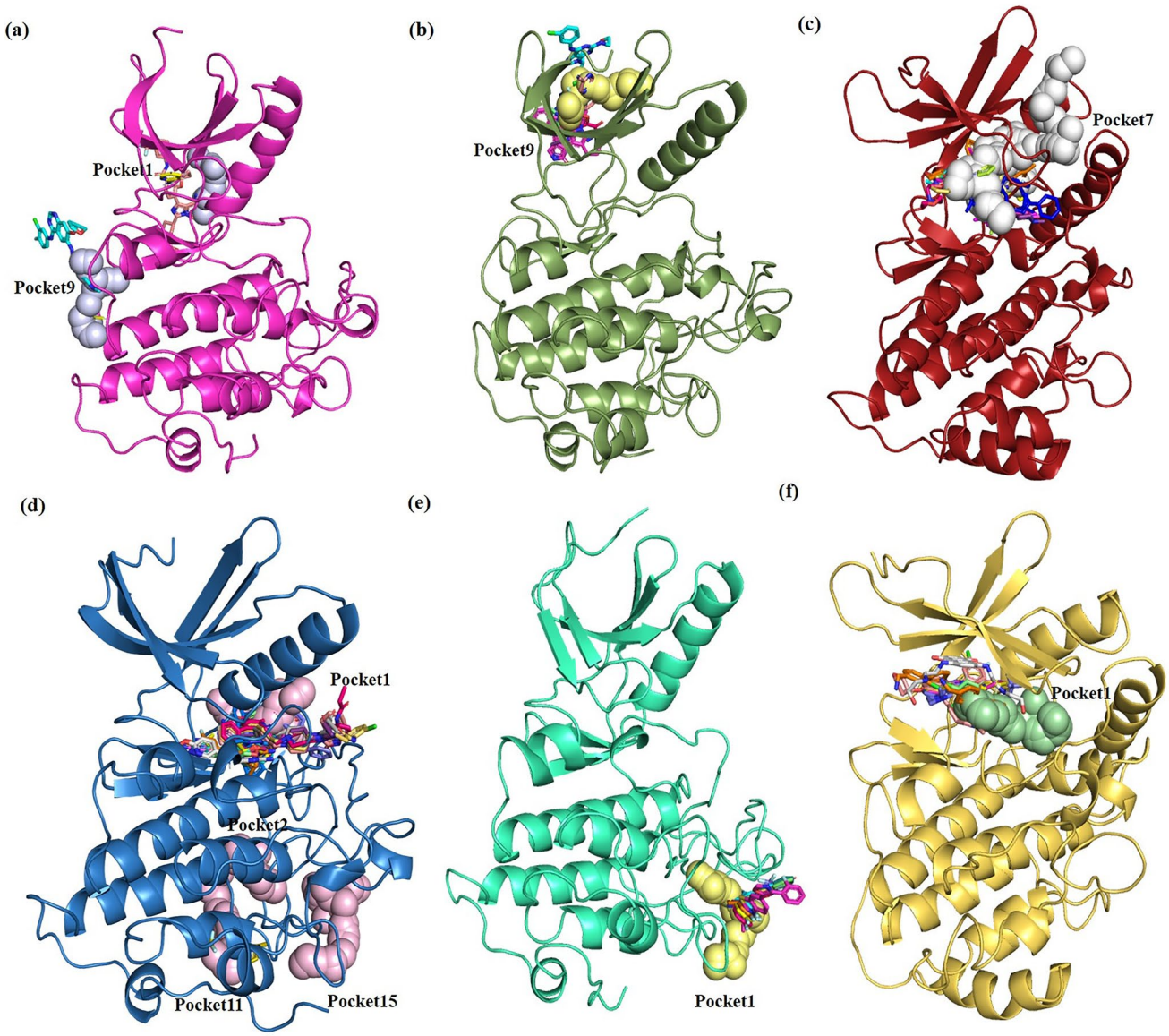

Fig. 9 Druggable pockets of the identified conformations and the overlaying respective orthosteric or allosteric ligands from KLIFS database, (**a**) CinterDI-CL (**b**) CinterDinter-CL (**c**) CODO-CL (**d**) CIDinter-ST (**e**) CODI-ST (**f**) CODO-ST

have optimal scaffold for accommodating well-known inhibitors in clinical trials Type2 inhibitors such as STI, NIL, AXI, BAX, B96, GV0, P30, P31, AV9, L1X, and Type3 inhibitors EUI, 573, 2AN, CCX, 1BR. Whilst in the P loop classification of the intermediate and inactive states, in the ST-based P.loop, the conformations CinterDI, CODinter, and CinterDinter did not feature optimal druggable pockets. Although the CODO-ST, CODI-ST, had one druggable pocket, the CODI-ST pocket is highly distal from the conserved kinase domain elements and did not show any overlap with any of the structural cavity bearing the drugs in clinical trials, while CODO-ST covers the linker region followed by αD and XDFG motif and was found to be overlapped with Type1 inhibitor, H8H, VH6 as this pocket. Amongst the available four druggable pockets of CIDinter, only Pocket11 was found to superimpose with Type1 inhibitor GD9 binding cavity.

Among the conformations with collapsed G/P.loop, CODI, and CinterDO didn't show any optimal druggable sites. The CIDO-CL spans a druggable pocket between αC and β4 and was observed to be significant overlap with the inhibitors in clinical trials such as Type2 inhibitors BAX, NIL, B96, P30, P31, L1X, and Type1.5 inhibitor (P06), and Type1 inhibitor (GD9, SD5). Pocket6 (spans around αC and b.l) and Pocket3 (comprises of linker and αD) of CinterDI-CL were found to accommodate Type1.5_Front

inhibitor (UM4) and Type1 inhibitor (0WM), respectively. In the case of CIDinter-CL, both the pockets did not overlap with any of the available orthosteric and allosteric ligands in the KLIFS database. Whereas, the CinterDinter-CL has a druggable pocket surrounded by the antiparallel β sheets of the N-lobe and was observed to be overlapped with 0WM (Type1), UM4 (Type1.5_Front), 573 (Type3) inhibitors. The CODO-CL druggable pocket was observed to have overlapped with the type of inhibitors as that of CODO-CT.

In comparison with the available co-crystal ligand (inhibitor bound) crystal structures of LIMK1 and LIMK2 (Fig. S18), all the co-crystal ligands are known to span at the N-lobe occupying the ATP binding site. Thereby among the identified druggable pockets, CIDI (Pocket1 & 2), CODO all the three classes of A.l and P.l namely CT (Pocket1 & 9), ST (Pocket1), CL (Pocket7), CIDO-CL (Pocket1) were observed to possess similar druggable pocket as that of the crystal structures. Moreover, the co-crystal ligands namely STU, 9DB, RXQ, and 35H available in other kinases have been overlapped with these pockets. However, the other pockets were not observed in the available LIMK1/2 crystal structures.

Conclusion

Protein kinases are one of the largest crucial drug targets in humans, however, the currently available chemical space for specific kinase inhibitor identification is highly limited. Hence, there is a need to identify an extended pharmacological space for accelerating the rational identification of conformation-specific inhibitors. In the current therapeutic research scenario, LIMK2 inhibitors are emerging as a potential therapeutic modality for treating various disorders, yet the conformational transitions from $DFG_{in}\alpha C_{in}$ (active) and $DFG_{out}\alpha C_{out}$ (inactive) states are unexplored. Therefore, in this study, the varying conformational states of LIMK2 were sampled to unravel the potential and novel druggable conformations through structural bioinformatics methods. Hence, in this study, molecular modelling followed by an extensive 1 microsecond (μs) timescale of molecular dynamics simulation was performed inclusive of perturbation response scanning and the community analysis towards the diverse conformational sampling of the kinase domain. This study deciphered the conformational transitions of LIMK2 that resulted in the following states, active (CIDI), intermediate/inactive states (CIDO, CinterDO, CinterDI, CODI, CIDinter, CinterDinter, CODinter) followed by inactive state (CODO). Through PRS map analysis, a maximum number of conserved kinase domain residues were observed to act as effectors only in CIDinter conformation among the intermediate/inactive states. Furthermore, it was also revealed that the conserved β-sheet, G/P.loop, and α helices (αC and αD) were observed as prominent effectors in most of the intermediate/inactive states thereby assisting in the conformational transitions among the states. In the case of structural network and metapath analysis, among the intermediate/inactive states, the CinterDinter conformation was observed to share a maximum number of nodes with the CIDI (active) state, whereas the CIDinter and CinterDO were the only intermediate states that had shared nodes with CODO (inactive) state. This implies that these corresponding intermediates are the highly successive and prior conformations of the CIDI and CODO states respectively, with the higher predominance of shared links and shared nodes. Furthermore, it was also inferred that the HRD motif and C.l residues act as the crucial nodes for the allosteric communications specifically, H449 a prominent node in both intermediates and CIDI conformation. Whilst, in CODO and intermediate states, the C.l and β3 were observed as prominent nodes that mediate the allosteric signalling, explicitly the N456 residue serves as a prominent node. Subsequent classification of the intermediate and inactives revealed that the intermediate states except for CIDO were observed to have A.l in the Open DFG-out state, thereby enabling the binding of the substrates and compatible with the binding of Type1 inhibitors. On contrary, the A.l in CODO (inactive) state was observed to be in a Closed type 2 state i.e., it unfavours the binding of substrate but highly accommodates Type2 inhibitors. In addition, the G/P.loop classification revealed that except for CIDO, CinterDO that has Collapsed G/P.loop, and CODinter with Stretched G/P.loop, all the other conformational states were observed to have both stretched and collapsed G/P.loop. Since studies have suggested that the collapsed G/P.loop has a higher predominance for inhibitor or substrate binding, except for CODinter all others are highly suitable for inhibitor or substrate binding. Finally, the druggable pocketome exploration revealed that except for CIDI and CODO state, all the other intermediates were observed to have druggable sites distal from ATP binding site, predominantly in the C-lobe that may act as an allosteric site for inhibitor identification. Moreover, among the classified states, the CODO and CIDI were observed to have prominent druggable pockets in terms of druggability score, volume, and overlay FDA-approved inhibitors namely Type1, Type1.5, Type2, and Type3. This mechanistic understanding of the large-conformational changes of LIMK2 will aid in the development of novel targeted therapeutics. In conclusion, this study provides potential insights into the intermediate conformational druggable states of LIMK2, based on this a series of inhibitors that have a distinct preference, especially for targeting the inactive states of LIMK2 shall be identified. To date, only the Type2 and Type3 inhibitors are available for targeting LIMK2 due to the lack of knowledge on its other potential conformational transitions. Thus, the druggable sites resultant from this study shall act as a

scaffold towards identifying ATP-competitive Type1, and Type1.5 inhibitors for LIMK2 through high throughput virtual screening methods and the identified hits for the respective states shall be validated by in vitro studies for the identification of the potential hits. In addition, all these insights shall collectively aid in identifying the druggable sites distal from ATP binding site for targeting LIMK2 in various disease conditions especially, in cancer and glaucoma.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10822-022-00459-0.

Acknowledgements JJ and NH greatly acknowledges the Department of Science and Technology, DST-FIST (SR/FST/LSI-667/2016), New Delhi for providing infrastructure facilities. JJ and NH sincerely acknowledge the Department of Science and Technology, New Delhi for the financial support in general and infrastructure facilities sponsored under the PURSE 2nd Phase programme (Order No. SR/ PURSE Phase 2/38 (G) dated: 21.02.2017), Department of Biotechnology, Bioinformatics Centre (BIC) -No.BT/PR40154/BTIS/137/34/2021. NH greatly acknowledges the Department of Science and Technology (DST), for providing financial assistance through DST/INSPIRE Fellowship/2019/IF190083. All authors would like to acknowledge Vision Research Foundation for hardware support.

Author contributions NH- Conceptualization, Investigation, Methodology, Validation, Writing—original draft; SA- Data curation, Investigation; UV- Supervision, Writing—review & editing; JJ- Investigation, Supervision, Writing—review & editing.

Declarations

Competing interests The authors declare no competing interests.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Roskoski R (2020) Properties of FDA-approved small molecule protein kinase inhibitors: a 2020 update. Pharmacol Res 152:104609. https://doi.org/10.1016/j.phrs.2019.104609
2. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. Science 298(5600):1912–1934. https://doi.org/10.1126/science.1075762
3. Kanev GK, de Graaf C, Westerman BA, de Esch IJP, Kooistra AJ (2021) KLIFS: an overhaul after the first 5 years of supporting kinase research. Nucleic Acids Res 49(D1):D562–D569. https://doi.org/10.1093/nar/gkaa895
4. Möbitz H (2015) The ABC of protein kinase conformations. Biochim Biophys Acta 1854:1555–1566. https://doi.org/10.1016/j.bbapap.2015.03.009
5. Roskoski R (2019) Properties of FDA-approved small molecule protein kinase inhibitors. Pharmacol Res 144:19–50. https://doi.org/10.1016/j.phrs.2019.03.006
6. Roskoski R (2021) Properties of FDA-approved small molecule protein kinase inhibitors: a 2021 update. Pharmacol Res 165:105463. https://doi.org/10.1016/j.phrs.2021.105463
7. Fabbro D, Cowan-Jacob SW, Moebitz H (2015) Ten things you should know about protein kinases: IUPHAR review 14. Br J Pharmacol 172(11):2675–2700. https://doi.org/10.1111/bph.13096
8. McClendon CL, Kornev AP, Gilson MK, Taylor SS (2014) Dynamic architecture of a protein kinase. Proc Natl Acad Sci USA 111(43):E4623–E4631. https://doi.org/10.1073/pnas.1418402111
9. Bhullar KS, Lagarón NO, McGowan EM, Parmar I, Jha A, Hubbard BP, Rupasinghe HPV (2018) Kinase-targeted cancer therapies: progress, challenges and future directions. Mol Cancer 17(1):48. https://doi.org/10.1186/s12943-018-0804-2
10. Ferguson FM, Gray NS (2018) Kinase inhibitors: the road ahead. Nat Rev Drug Discov 17(5):353–377. https://doi.org/10.1038/nrd.2018.21
11. Manetti F (2012) LIM kinases are attractive targets with many macromolecular partners and only a few small molecule regulators. Med Res Rev 32(5):968–998. https://doi.org/10.1002/med.20230
12. Sumi T, Matsumoto K, Nakamura T (2001) Specific activation of LIM kinase 2 via phosphorylation of threonine 505 by ROCK, a Rho-dependent protein kinase. J Biol Chem 276(1):670–676. https://doi.org/10.1074/jbc.M007074200
13. Bamburg JR (1999) Proteins of the ADF/cofilin family: essential regulators of actin dynamics. Annu Rev Cell Dev Biol 15:185–230. https://doi.org/10.1146/annurev.cellbio.15.1.185
14. Scott RW, Olson MF (2007) LIM kinases: function, regulation and association with human disease. J Mol Med (Berl) 85(6):555–568. https://doi.org/10.1007/s00109-007-0165-6
15. Harrison BA, Whitlock NA, Voronkov MV, Almstead ZY, Gu K-j, Mabon R, Gardyan M, Hamman BD, Allen J, Gopinathan S, McKnight B, Crist M, Zhang Y, Liu Y, Courtney LF, Key B, Zhou J, Patel N, Yates PW, Liu Q, Wilson AGE, Kimball SD, Crosson CE, Rice DS, Rawlins DB (2009) Novel class of LIM-kinase 2 inhibitors for the treatment of ocular hypertension and associated glaucoma. J Med Chem 52(21):6515–6518. https://doi.org/10.1021/jm901226j
16. Manetti F (2018) Recent advances in the rational design and development of LIM kinase inhibitors are not enough to enter clinical trials. Eur J Med Chem 155:445–458. https://doi.org/10.1016/j.ejmech.2018.06.016
17. Shen M, Zhou S, Li Y, Li D, Hou T (2013) Theoretical study on the interaction of pyrrolopyrimidine derivatives as LIMK2 inhibitors: insight into structure-based inhibitor design. Mol Biosyst 9(10):2435–2446. https://doi.org/10.1039/c3mb70168a
18. Modi V, Dunbrack RL (2019) Defining a new nomenclature for the structures of active and inactive kinases. Proc Natl Acad Sci USA 116(14):6818–6827. https://doi.org/10.1073/pnas.1814279116
19. Ung PM-U, Rahman R, Schlessinger A (2018) Redefining the protein kinase conformational space with machine learning. Cell Chem Biol 25(7):916-924.e2. https://doi.org/10.1016/j.chembiol.2018.05.002
20. Goodwin NC, Cianchetta G, Burgoon HA, Healy J, Mabon R, Strobel ED, Allen J, Wang S, Hamman BD, Rawlins DB (2015) Discovery of a type III inhibitor of LIM kinase 2 that binds in a DFG-out conformation. ACS Med Chem Lett 6(1):53–57. https://doi.org/10.1021/ml500242y
21. Mathea S, Salah E, Hanke T, Newman JA, Oerum S, Wang D, Burgess-Brown N, Delft F von, Arrowsmith CH, Edwards AM, Bountra C, Bullock AN, Knapp S (2017) LIM Domain Kinase 2 (LIMK2) In Complex With TH-300. https://doi.org/10.2210/pdb5NXD/pdb
22. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U, Sali A (2006) Comparative protein structure modeling using modeller. Curr Protoc Bioinform. https://doi.org/10.1002/0471250953.bi0506s15

23. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18(12):1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12%3c1463:AID-JCC4%3e3.0.CO;2-H
24. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N ·log(N) method for Ewald sums in large systems. J Chem Phys 98(12):10089–10092. https://doi.org/10.1063/1.464397
25. Caliandro R, Rossetti G, Carloni P (2012) Local fluctuations and conformational transitions in proteins. J Chem Theory Comput 8(11):4775–4785. https://doi.org/10.1021/ct300610y
26. Atilgan C, Atilgan AR (2009) Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. PLoS Comput Biol 5(10):e1000544. https://doi.org/10.1371/journal.pcbi.1000544
27. Bakan A, Meireles LM, Bahar I (2011) ProDy: protein dynamics inferred from theory and experiments. Bioinformatics 27(11):1575–1577. https://doi.org/10.1093/bioinformatics/btr168
28. Bakan A, Dutta A, Mao W, Liu Y, Chennubhotla C, Lezon TR, Bahar I (2014) Evol and ProDy for bridging protein sequence evolution and structural dynamics. Bioinformatics 30(18):2681–2683. https://doi.org/10.1093/bioinformatics/btu336
29. Dutta A, Krieger J, Lee JY, Garcia-Nafria J, Greger IH, Bahar I (2015) Cooperative dynamics of intact AMPA and NMDA glutamate receptors: similarities and subfamily-specific differences. Structure 23(9):1692–1704. https://doi.org/10.1016/j.str.2015.07.002
30. General IJ, Liu Y, Blackburn ME, Mao W, Gierasch LM, Bahar I (2014) ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones. PLoS Comput Biol 10(5):e1003624. https://doi.org/10.1371/journal.pcbi.1003624
31. Astl L, Verkhivker GM (2019) Atomistic modeling of the ABL kinase regulation by allosteric modulators using structural perturbation analysis and community-based network reconstruction of allosteric communications. J Chem Theor Comput 15(5):3362–3380. https://doi.org/10.1021/acs.jctc.9b00119
32. Vijayabaskar MS, Vishveshwara S (2010) Interaction energy based protein structure networks. Biophys J 99(11):3704–3715. https://doi.org/10.1016/j.bpj.2010.08.079
33. Bowerman S, Wereszczynski J (2016) Detecting allosteric networks using molecular dynamics simulation. Methods Enzymol 578:429–447. https://doi.org/10.1016/bs.mie.2016.05.027
34. James KA, Verkhivker GM (2014) Structure-based network analysis of activation mechanisms in the ErbB family of receptor tyrosine kinases: The regulatory spine residues are global mediators of structural stability and allosteric interactions. PLoS ONE 9(11):e113488. https://doi.org/10.1371/journal.pone.0113488
35. Qiu Y, Yin X, Li X, Wang Y, Fu Q, Huang R, Lu S (2021) Untangling dual-targeting therapeutic mechanism of epidermal growth factor receptor (EGFR) based on reversed allosteric communication. Pharmaceutics 13(5):747. https://doi.org/10.3390/pharmaceutics13050747
36. Felline A, Seeber M, Fanelli F (2020) webPSN v2.0: A webserver to infer fingerprints of structural communication in biomacromolecules. Nucleic Acids Res 48(W1):W94–W103. https://doi.org/10.1093/nar/gkaa397
37. Schwarz D, Merget B, Deane C, Fulle S (2019) Modeling conformational flexibility of kinases in inactive states. Proteins 87(11):943–951. https://doi.org/10.1002/prot.25756
38. Steinberg SF (2018) Post-translational modifications at the ATP-positioning G-loop that regulate protein kinase activity. Pharmacol Res 135:181–187. https://doi.org/10.1016/j.phrs.2018.07.009
39. Le Guilloux V, Schmidtke P, Tuffery P (2009) Fpocket: an open source platform for ligand pocket detection. BMC Bioinformatics 10(1):99. https://doi.org/10.1186/1471-2105-10-168
40. Kaynak BT, Bahar I, Doruker P (2020) Essential site scanning analysis: a new approach for detecting sites that modulate the dispersion of protein global motions. Comput Struct Biotechnol J 18(5):1577–1586. https://doi.org/10.1016/j.csbj.2020.06.020
41. Kooistra AJ, Kanev GK, van Linden OPJ, Leurs R, de Esch IJP, de Graaf C (2016) KLIFS: a structural kinase-ligand interaction database. Nucleic Acids Res 44(D1):D365–D371. https://doi.org/10.1093/nar/gkv1082
42. van Linden OPJ, Kooistra AJ, Leurs R, de Esch IJP, de Graaf C (2014) KLIFS: a knowledge-based structural database to navigate kinase-ligand interaction space. J Med Chem 57(2):249–277. https://doi.org/10.1021/jm400378w
43. Bournez C, Carles F, Peyrat G, Aci-Sèche S, Bourg S, Meyer C, Bonnet P (2020) Comparative assessment of protein kinase inhibitors in public databases and in PKIDB. Molecules 25(14):3226. https://doi.org/10.3390/molecules25143226
44. Carles F, Bourg S, Meyer C, Bonnet P (2018) PKIDB: a curated, annotated and updated database of protein kinase inhibitors in clinical trials. Molecules 23(4):908. https://doi.org/10.3390/molecules23040908
45. Collie GW, Michaelides IN, Embrey K, Stubbs CJ, Börjesson U, Dale IL, Snijder A, Barlind L, Song K, Khurana P, Phillips C, Storer RI (2021) Structural basis for targeting the folded P-loop conformation of c-MET. ACS Med Chem Lett 12(1):162–167. https://doi.org/10.1021/acsmedchemlett.0c00392