



## OPEN

## Structural dynamic analysis of apo and ATP-bound IRAK4 kinase

## SUBJECT AREAS:

COMPUTATIONAL  
MODELSPROTEIN STRUCTURE  
PREDICTIONS

Vijayakumar Gosu &amp; Sangdun Choi

Department of Molecular Science and Technology, Ajou University, Suwon 443-749, Korea.

Received  
28 January 2014Accepted  
2 July 2014Published  
18 July 2014Correspondence and  
requests for materials  
should be addressed to  
S.C. (sangdunchoi@  
ajou.ac.kr)

Interleukin-1 receptor-associated kinases (IRAKs) are Ser/Thr protein kinases that play an important role as signaling mediators in the signal transduction facilitated by the Toll-like receptor (TLR) and interleukin-1 receptor families. Among IRAK family members, IRAK4 is one of the drug targets for diseases related to the TLR and IL-1R signaling pathways. Experimental evidence suggests that the IRAK4 kinase domain is phosphorylated in its activation loop at T342, T345, and S346 in the fully activated state. However, the molecular interactions of subdomains within the active and inactive IRAK4 kinase domain are poorly understood. Hence, we employed a long-range molecular dynamics (MD) simulation to compare apo IRAK4 kinase domains (phosphorylated and unphosphorylated) and ATP-bound phosphorylated IRAK4 kinase domains. The MD results strongly suggested that lobe uncoupling occurs in apo unphosphorylated IRAK4 kinase via the disruption of the R334/T345 and R310/T345 interaction. In addition, apo unphosphorylated trajectory result in high mobility, particularly in the N lobe, activation segment, helix  $\alpha$ G, and its adjoining loops. The Asp-Phe-Gly (DFG) and His-Arg-Asp (HRD) conserved kinase motif analysis showed the importance of these motifs in IRAK4 kinase activation. This study provides important information on the structural dynamics of IRAK4 kinase, which will aid in inhibitor development.

Toll-like receptors (TLRs) respond to pathogen-associated molecular patterns (PAMPs) and initiate the first wave of inflammatory signals and innate immune responses<sup>1,2</sup>. Interleukin-1 receptor (IL-1R) family cytokine receptors originate and regulate inflammatory and immune responses. The uncontrolled TLR/IL-1R signaling pathways, which involve interleukin-1 receptor-associated kinase 4 (IRAK4), lead to multiple diverse diseases ranging from chronic to autoinflammatory disorders<sup>3,4</sup>. Therefore, the upstream proteins involved in the activation of IRAK4 and signaling can be potential therapeutic drug targets. Several attempts have been made to modulate the activity of the drug target, and some of the drugs that were designed for these targets are now in clinical trials<sup>4,5</sup>. More than a few novel classes of inhibitors have been developed recently for the inhibition of IRAK4 activation. However, these inhibitors are still in preclinical trials<sup>6</sup>.

IRAKs (Ser/Thr kinases) are essential mediators in TLR and IL-1R signal transduction<sup>7,8</sup>. Upon stimulation, TLRs and IL-1Rs recruit the adaptor proteins and thereby form an intracellular receptor-adaptor complex<sup>9,10</sup>. This receptor-adaptor complex in turn recruits IRAKs<sup>11</sup>. Upon IRAK recruitment, IRAK1 undergoes phosphorylation by IRAK4 on key Ser and Thr residues. IRAK1 then detaches from the receptor complex and associates with tumor necrosis factor receptor-associated factor 6 (TRAF6); then, it binds with the TGF- $\beta$ -activated kinase 1-TAK1-binding protein 1-TAK1-binding protein 2 (TAK1-TAB1-TAB2) complex, leading to the stimulation of nuclear factor  $\kappa$ B (NF- $\kappa$ B), p38, and c-Jun N-terminal protein kinase mitogen-activated protein kinases (JNK MAPKs)<sup>12,13</sup>. Several studies have shown the involvement of IRAK4 in TLR and IL-1R signaling, particularly in Myd88-dependent TLR signaling<sup>14</sup>. Severe defects have been observed in cytokine responses in mice lacking IRAK4, and impaired responses have been seen in human IRAK4-deficient cells<sup>14-16</sup>. These studies indicate that, despite the absolute involvement of IRAK4 in TLR signaling, inhibition of IRAK4 activation may cause too broad an impact on TLR/IL-1R signaling pathways. Deficiency of IRAK4 in children makes them susceptible to streptococcal pyogenic infection, but no severe viral or parasitic infection has been observed<sup>17</sup>. In IRAK4-deficient patients approaching adolescence, susceptibility to infections, particularly chronic infections, decreases<sup>18,19</sup>.

IRAK4 may induce signals via intermolecular interactions and its kinase activity, resulting in activation of the downstream kinase IRAK1<sup>14</sup>. Experimental studies have reported that some pathways, such as NF- $\kappa$ B and JNK, require IRAK4 kinase function. However, cells expressing only inactive mutants of IRAK4 show reduced response in NF- $\kappa$ B activation compared to the wild type<sup>20</sup>. Studies on IRAK4 kinase-dead mutant knock-in mice from different strains also support the significance of IRAK4 kinase activity<sup>21-24</sup>. Another report has also indicated that disruption of the formation of the receptor complex with a truncated kinase domain of IRAK4 inhibits IL-1 signaling<sup>17</sup>. TLRs, such as TLR7 and TLR9, activate IRFs, which induce type I interferon (IFN). IRAK4-dependent



induction of IFNs has been observed for TLR7 and TLR9 signaling upon stimulation with TLR ligands<sup>25</sup>. The amelioration of disease symptoms observed in lupus-prone mice upon dual inhibition of TLR7 and TLR9 suggests that IRAK4 may be a potential therapeutic drug target for systemic lupus erythematosus (SLE)<sup>26,27</sup>. A recent study has also reported that IRAK4 and IRAK1 are associated with Vogt-Koyanagi-Harada (VKH) disease<sup>28</sup>. Therefore, therapeutic benefits can be achieved by modulating the function of IRAK4 via kinase inhibition for toning down the inflammatory responses in TLR/IL-1R signaling.

The mammalian IRAK family comprises four members: IRAK1<sup>29</sup>, IRAK2<sup>30</sup>, IRAKM<sup>31</sup> and IRAK4<sup>14</sup>. All IRAK family members show similar domain structure with an N-terminal death domain and a central kinase domain (KD). IRAK4 shares high similarity in the KD with all other family members<sup>8</sup>. The crystal structure of IRAK4 KD reveals that it is composed of two lobes: an N-terminal lobe comprising mainly antiparallel  $\beta$ -strands and an  $\alpha$ -helix (helix  $\alpha$ C); and a C-terminal lobe exclusively comprising  $\alpha$ -helices (Figure 1). The hinge sequence connecting N and C lobes partially describes the adenosine triphosphate (ATP) and the ATP-competitive kinase inhibitors<sup>32</sup>. The N lobe contains the glycine-rich loop (GXGX $\phi$ G) formed by the first two  $\beta$ -strands. The activation segment (329–358) located in the C lobe contains important catalytic residues such as the Asp-Phe-Gly (DFG) (329–331) motif. The central region of the activation segment is the activation loop; autophosphorylation of this activation loop at T342, T345, and S346 is required for IRAK4 kinase activity, and mutation at these positions causes significant reduction in the catalytic activity of IRAK4 kinase<sup>32–34</sup>. Out of these three phosphorylation sites, T345 is the only prototypical phosphoresidue required for the activity of IRAK4 kinase, which is similar to many protein kinases<sup>35,36</sup>. In IRAK4, pT345 exhibits direct interaction with R334 and with the S346 backbone amine. In addition, phosphorylation at T342 (pT342) is coordinated by T365, K367, and K441; phosphorylation at S346 (pS346) is highly exposed to the solvent, without any intramolecular interactions<sup>32,34</sup>. Structural ana-

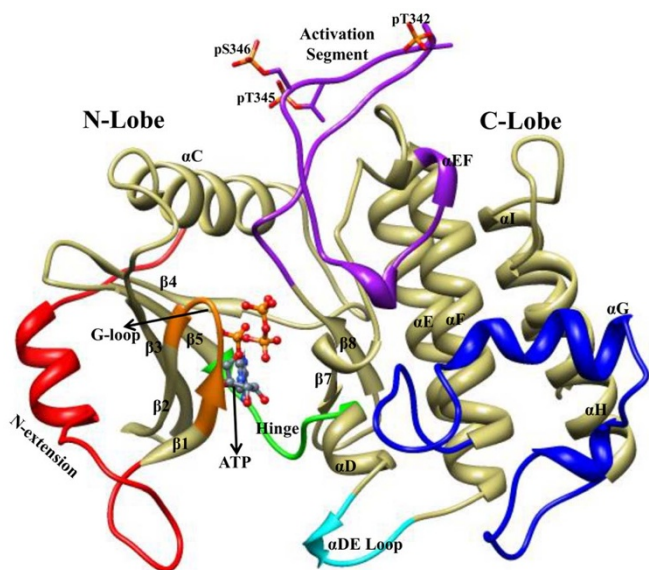
lysis of inactive kinases has revealed that the unphosphorylated activation segment can adopt different conformations<sup>37,38</sup>. The His-Arg-Asp (HRD) (309–311) motif is important in the catalytic loop and Arg from the HRD motif is important for the kinases regulated by the activation loop<sup>35</sup>. The apo-active crystal structure of IRAK4 KD has been found to have two distinct conformations for helix  $\alpha$ C in the N lobe: helix  $\alpha$ C-in and helix  $\alpha$ C-out. Although the helix  $\alpha$ C-in is the active conformation, the usual salt bridge between Lys (K213) and Glu (E233) is absent, in contrast to the finding for other kinases<sup>39,40</sup>. Most importantly, IRAK4 has a Tyr (Y262) as the gatekeeper residue, which forms a hydrogen bond with E233 in the active form of the kinase; thus, the back pocket is not accessible to ligands. The other IRAK family members IRAK1, IRAK2 and IRAKM also have Tyr as a gatekeeper residue<sup>32</sup>.

Over the past few years, molecular dynamics (MD) simulation has been extensively employed as a computational approach for investigating the connecting pathway between the active and inactive states of kinase. For example, the transition from the inactivation state to the activation state of the Src tyrosine kinase catalytic domain was studied by the string method with swarms of trajectories. This study suggested that activation occurs via the activation loop first, followed by helix  $\alpha$ C rotation<sup>41</sup>. A metadynamics sampling approach used to study the mechanism underlying the transition from the activation to inactivation state of cyclin-dependent kinase 5 (CDK5) suggests a two-step mechanism, that is, helix  $\alpha$ C rotates by  $\sim 45^\circ$ , which is followed by activation loop refolding<sup>42</sup>. CDK2 was studied by Groban *et al.* using molecular mechanics to reproduce the conformational changes that take place on CDK2 phosphorylation. This study suggests that changes occur in the energy landscape because of phosphorylation<sup>43</sup>. Hematopoietic cell kinase (HCK) was studied using targeted MD simulations to investigate the transition from the inactive to the active conformation. The results of this study suggest the structural effects due to phosphorylation of the activation loop are sequentially transmitted to the helix  $\alpha$ C and the hinge region<sup>44</sup>. PrkC kinase from *Bacillus subtilis* were studied using MD simulations to examine the conformational changes caused by phosphorylation and ATP binding<sup>45</sup>. In addition, catalytically inactive EGFR conformations of an active EGFR kinase were recently observed<sup>46</sup>.

Since there is very little experimental evidence available for IRAK4 kinase, the regulation of the mechanism of activation is not clearly understood. However, in comparison with other kinases, regulation of IRAK4 kinase may depend on activation loop phosphorylation and switching between helix  $\alpha$ C-in and helix  $\alpha$ C-out conformations<sup>34</sup>. Therefore, the activation and inactivation cycle of IRAK4 must be a multi-component process with many levels of regulation. To date, no crystal coordinates are available for inactive IRAK4 kinase. Recently, IRAK4 kinase has attracted attention as a therapeutic drug target in TLR signaling<sup>6</sup>. Hence, it is necessary to understand the inactive IRAK4 kinase to unravel the regulation mechanism of IRAK4 kinase. In this study, we have applied computational approaches to compare apo IRAK4 KDs (phosphorylated and unphosphorylated) and ATP-bound phosphorylated IRAK4 KDs. For simplicity, hereafter we designate unphosphorylated apo IRAK4 KD as uIRAK4, phosphorylated apo IRAK4 KD as pIRAK4, and ATP-bound phosphorylated IRAK4 KD as pIRAK4-ATP whenever necessary. Most importantly, we carried out 200-ns long-range MD simulation for all the three forms of IRAK4 KDs to understand the conformational changes in uIRAK4 and pIRAK4, and also examined the impact of the natural ligand ATP on the structural changes in IRAK4 KD. In addition, principal component analysis (PCA) was performed to understand the collective motions within uIRAK4, pIRAK4, and pIRAK4-ATP.

## Results

**Summary of the IRAK4 KD systems.** To compare and understand the conformational changes of apo and ATP-bound forms of IRAK4



**Figure 1 | Structure of IRAK4 kinase domain.** The structure of IRAK4 KD predominantly shows the N lobe with  $\beta$ -strands and an  $\alpha$ -helix, C lobe with exclusively  $\alpha$ -helices. The structural parts were shown in various colors, N-terminal extension preceding kinase domain in red, G loop in orange, activation segment in magenta, hinge region in green,  $\alpha$ DE loop in cyan, helix  $\alpha$ G and its adjoining loops in blue. ATP molecule between the N and C lobe were shown in ball and stick model in slate gray color. Phosphorylated residues pT342, pT345, and pS346 are highlighted. The structural parts were labeled according to the crystal structure<sup>32</sup>.



Table 1 | Summary of the trajectories subjected to molecular dynamics simulations

Trajectory	Phosphorylation state	Ligand	Components
uIRAK4	Unphosphorylated	Apo	uIRAK4 in water
pIRAK4	pT342, pT345, and pS346	Apo	pIRAK4 in water
pIRAK4-ATP	pT342, pT345, and pS346	ATP and 2 Mg <sup>2+</sup> ions	pIRAK4, ATP, and 2 Mg <sup>2+</sup> ions in water

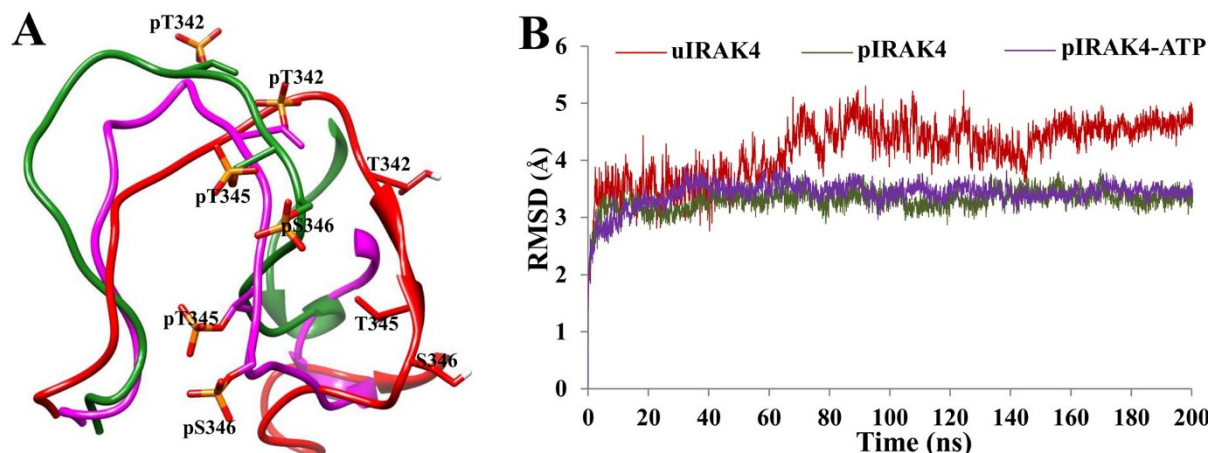
KD, the structural models were constructed for apo and ATP-bound IRAK4 KDs, as detailed in Table 1. The structural models uIRAK4, pIRAK4, and pIRAK4-ATP were subjected to dynamics simulations for the production run of 200 ns (refer to the methods section).

**Conformational changes in the activation segment.** It is well known that phosphorylation of the activation loop is required for the activation of many kinases<sup>47</sup> and that the activation loop undergoes conformational changes upon switching between inactive and active states of the kinase. The number and location of phosphorylation sites within the activation loop varies between kinases<sup>48</sup>. T342, T345, and S346 are the phosphorylation sites in the activation loop of IRAK4 KD<sup>33</sup>. We were interested in understanding the structural changes in the activation segment (329–358) of the IRAK4 systems due to the phosphorylation of the activation loop. To investigate this, we used the SuperPose server to superimpose the activation segments from uIRAK4, pIRAK4, and pIRAK4-ATP obtained from the MD simulation (Figure 2A). The RMSD values of the activation segments of uIRAK4 and pIRAK4, uIRAK4 and pIRAK4-ATP, and pIRAK4 and pIRAK4-ATP were 4.78 Å, 5.22 Å, and 4.15 Å, respectively. In particular, the RMSD between pIRAK4 and pIRAK4-ATP was lower than the values of uIRAK4 and pIRAK4 and of uIRAK4 and pIRAK4-ATP. These RMSD values indicate a large structural variation in the activation segment upon activation loop phosphorylation and ATP binding to the phosphorylated IRAK4 kinase in order to stabilize the activation segment. To further elucidate the structural flexibility in the activation segment, we superimposed ten snapshots from 200-ns MD run (Supplementary Information Figure S1). From these results, we observed that the activation segment remained relatively stable in pIRAK4-ATP as expected. In contrast, both phosphorylated and unphosphorylated structures were unstable. The conformational flexibility in the pIRAK4 suggests that phosphorylation in the activation loop induces conformational changes in the activation segment. In addition, deviations in the activation segments upon removal of phosphate groups (uIRAK4)

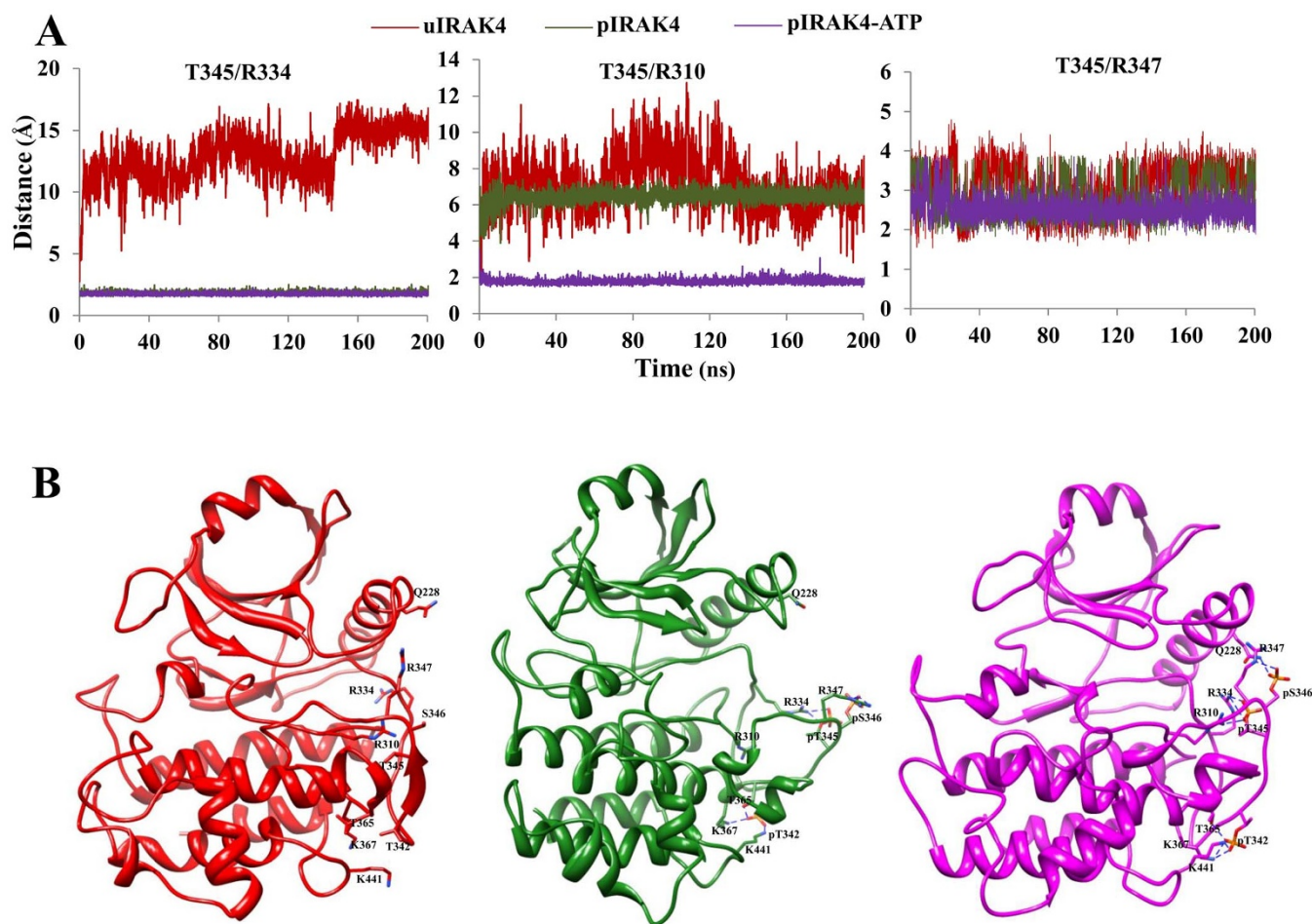
indicated that the initial inactivation dynamics of IRAK4 KD may impose some conformational changes on the activation segment.

To obtain more in-depth understanding of the structural plasticity in the activation segment, we focused on the G331 from the DFG (329–331) motif and T345 in the activation loop, as these residues are important for kinase activity. The DFG motif is highly conserved in eukaryotic Ser/Thr kinase and is situated at the base of the activation loop. In the activation state of the kinase, ATP forms polar interactions with DFG, either directly or indirectly, via magnesium ions. Distortions in the DFG motif have also been observed in some kinases due to the structural changes that originate from the activation loop phosphorylation<sup>49</sup>. The prototypical residue pT345 in the activation loop coordinates with the basic amino acids important for kinase activity. G331 from DFG exhibits vertical and horizontal motions in both pIRAK4 and uIRAK4. In the pIRAK4-ATP model, the G331 position is highly stable. This suggests that activation loop phosphorylation, or unphosphorylation causes the distortions in the DFG loop. The position of T345 in the pIRAK4-ATP model differed from that in the pIRAK4 and uIRAK4 models (Supplementary Information Figure S1). In pIRAK4 and uIRAK4, the position of the T345 showed horizontal and vertical motion, respectively, whereas in the pIRAK4-ATP model, the position of pT345 was stable with vertical motion (Supplementary Information Figure S1). In particular, in pIRAK4 and pIRAK4-ATP, the phosphate group of pT345 was located close to the basic amino acids necessary for the stability of its activation segment, thereby altering the shape of the activation loop. In contrast, T345 in the activation loop of uIRAK4 was located away from the basic amino acids.

The RMSD plot for the residues of activation segment (Figure 2B) showed that uIRAK4 has larger deviations compared with pIRAK4 and pIRAK4-ATP. It indicated that the phosphorylation of the activation loop and ATP binding to IRAK4 kinase may stabilize and lead to conformational changes in the activation segment, and thus may activate the kinase and allow for substrate access<sup>36</sup>. The RMSD of the DFG motif can also be correlated with the above results,



**Figure 2 | Superimposition and RMSD of the activation segment.** (A) Activation segment superimposition of apo and ATP-bound structures. (B) Residue RMSD of the activation segments with respect to the initial structure. Apo unphosphorylated in red, apo phosphorylated in forest green, and phosphorylated with bound ATP in magenta color are given. These color codes are used for all the figures to represent apo and ATP-bound structures.



**Figure 3 | Molecular interaction of T345.** (A) Residue minimum distances of T345/R334, T345/R310, and T345/R347. (B) Apo and ATP-bound IRAK4 KD is shown in ribbon model. Phosphorylation sites and interacting residues are shown in stick model. The hydrogen bond and salt bridge interactions are shown with blue dashes. The color codes are given in the Figure 2 legend.

showing that the pIRAK4-ATP structure is highly stable after  $\sim 35$  ns compared with both pIRAK4 and uIRAK4 structures (Supplementary Information Figure S2A). Additionally, the RMSF of the activation segment shows that there is high residue flexibility in uIRAK4 in the T345 position compared with pIRAK4 and pIRAK4-ATP (Supplementary Information Figure S2B). The above results cumulatively suggest that even though the phosphorylation of the activation loop induces conformational changes in the DFG motif, activation loop phosphorylation and ATP binding are mandatory for the stabilization of the activation loop, and thus the whole activation segment is stabilized<sup>33,45</sup>.

**Stabilization of the basic residues (Arg triad) at the phosphoT345 site.** In general, the critical electrostatic interaction of a specific phosphoresidue in the activation loop with a positively charged (Arg) cluster leads to activation loop refold, and the loop itself is positioned for substrate binding. The active crystal structure of IRAK4 KD revealed that its activation loop phosphorylation is similar to that of other phosphorylated Ser/Thr/Tyr kinases<sup>32,34</sup>. However, out of the three phosphorylation sites pT342, pT345, and pS346 in the activation loop, T345 is the primary phosphoresidue anchor toward the basic pocket that is conserved among kinases<sup>50,51</sup>. R310 precedes the catalytic base, and R334 and R347 in the activation loop are coordinated with the primary phosphoresidue pT345. Out of the three basic residues coordinating with the phosphoresidue pT345, the active crystal structure revealed that R334 has direct interaction with the

phosphate group of T345. To check the consistency of the interactions between the phosphate group of pT345 and the basic pocket, we examined the MD trajectories of uIRAK4, pIRAK4, and pIRAK4-ATP. MD simulations of pIRAK4 and pIRAK4-ATP revealed that R334 is coordinated with the phosphate group of pT345, with minimum residue distance around 2.5 Å throughout the simulations. In contrast, on removal of the phosphate group in uIRAK4, we observed that the residue minimum distance of T345/R334 consistently increases to  $\sim 17$  Å (Figure 3A). Interestingly, we also noticed that the minimum distance of pT345/R310 during the simulations was within 2.5 Å in pIRAK4-ATP, whereas this distance was up to 7 Å and  $\sim 8$  Å (till  $\sim 120$  ns, distance raised up to 10 Å in uIRAK4) in pIRAK4 and uIRAK4 structures, respectively (Figure 3A). In addition, the minimum distance of pT345/R347 was 2–4 Å in all three trajectories, with high fluctuation during simulations (Figure 3A). Firstly, the minimum distances of T345/R310 and T345/R334 observed in all three simulations suggest that R310 from the catalytic loop and R334 from the activation loop play dominant roles in the anchoring of pT345, compared with R347 upon ATP binding to the phosphorylated IRAK4 kinase. Secondly, there is an indication that interaction between the N lobe and C lobe loosens in inactive IRAK4 kinase through the disruption of the interaction of T345/R310 and T345/R334. The uncoupled T345/R310 and T345/R334 fluctuations may reflect the intrinsic dynamics in IRAK4 kinase<sup>32,34</sup>. These results may also indicate that R310 and R334 complement each other upon natural ligand ATP binding. The minimum distance of T345/R310 in pIRAK4, that is,



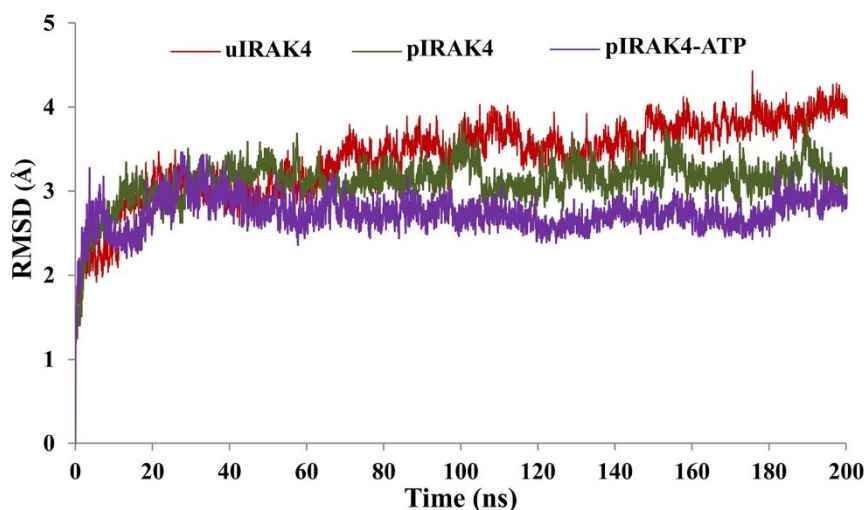
7 Å, and of 2.5 Å for pIRAK4-ATP during simulation further indicates the role of R310 in the transition from the activation to the inactivation state by loss of interaction between the phosphate group of pT345 and the side chain of R310. It is to be noted that this type of interaction is important for most of the RD kinases<sup>36,48</sup>.

**Molecular interactions at phosphoT342 and S346 sites.** To explore the importance of pT342 and pS346 upon activation of IRAK4 kinase, we analyzed the interaction pattern of pT342 and pS346 with their surrounding residues during simulations. The residue pT342 forms stable electrostatic interaction with T365, K367, and K441 in both pIRAK4 and pIRAK4-ATP structural models, which is consistent with the crystal coordinates. Interestingly, we observed that pS346 shows hydrogen bond interaction with Q228 from helix  $\alpha$ C in the pIRAK4-ATP model (Figure 3B). Q228 from helix  $\alpha$ C is quite distant from S346 in the uIRAK4 and pIRAK4 model, that is  $\sim 18$  Å and  $\sim 15$  Å, respectively (Supplementary Information Figure S3). In the active crystal structure with AMP-PNP binding to IRAK4 KD, pS346 is highly exposed to solvents without any intramolecular interaction. However, in our analysis, upon ATP and Mg<sup>2+</sup> ion binding to IRAK4 KD (pIRAK4-ATP), the side chain of the Q228 residue reorients itself toward pS346, suggesting that, in addition to the Arg cluster, Q228 might also play an important role in switching the helix  $\alpha$ C to the proper orientation for the catalytic residue E233. In most of the Ser/Thr kinases, the basic residues in the helix  $\alpha$ C interact with phosphorylated residue in the activation loop<sup>52</sup>. In addition, upstream of pS346, the side chain of R347 also reorients itself toward pS346 upon phosphorylation and ATP binding. Mutation of this phosphorylation site reduces the catalytic activity of IRAK4 kinase up to 50%<sup>33</sup>, reflecting the importance of this phosphorylation site in the regulation mechanism of IRAK4 kinase and indicating that this phosphorylation site of IRAK4 KD warrants further investigation.

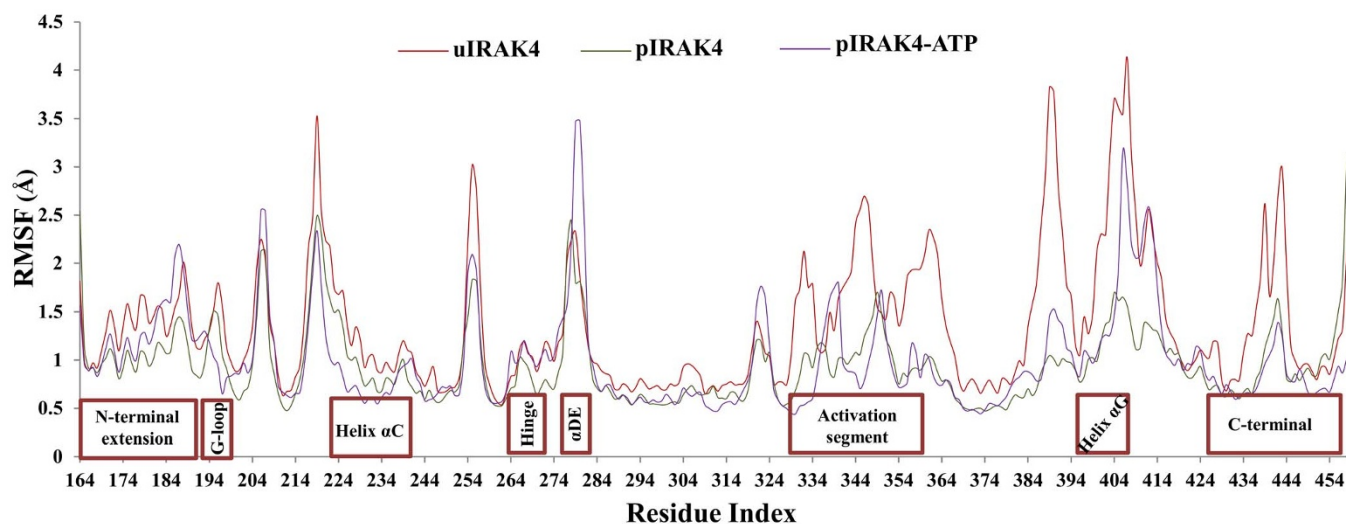
**Deviations from initial structure and residue flexibility analysis.** RMSD values were used to investigate the structural difference between the apo and ATP-bound forms of IRAK4 KD. RMSD values of the backbone atoms for uIRAK4, pIRAK4, and pIRAK4-ATP are plotted as a function of time in Figure 4. The RMSD plot shows that all the three trajectories were able to reach reasonable equilibrium at  $\sim 28$  ns. As expected, the ATP-bound phosphorylated structure (pIRAK4-ATP) is more stable than the apo trajectories, indicating that ATP induces conformational changes in order to stabilize the whole structure, whereas both the apo trajectories

uIRAK4 and pIRAK4 show slightly higher fluctuations, even after equilibration. In particular, uIRAK4 induces rapid deviation even after  $\sim 60$  ns. It is noteworthy that pIRAK4 shows slight stable deviation, indicating that phosphorylation of the activation loop stabilizes the activation segment, thereby altering the whole conformation of the structure followed by ATP binding. To identify the most flexible regions of apo and ATP-bound IRAK4 KD, the RMSFs of the backbone atoms of each residue were calculated over the last 100 ns (100–200 ns) and are plotted in Figure 5. The flexible regions of both the apo trajectories are the G loop, the loop formed by the 2<sup>nd</sup> and 3<sup>rd</sup>  $\beta$ -strand, the loop at the N terminus of the helix  $\alpha$ C, the N terminus of the hinge region in the N lobe, the  $\alpha$ DE loop (276–282), the activation segment (329–358) (only for uIRAK4), the helix  $\alpha$ EF, and the helix  $\alpha$ G (395–406) (only for uIRAK4) and its adjoining loops (383–394 and 407–423) in the C lobe. Most of the Ser/Thr kinases show that these regions are involved in ATP and substrate binding<sup>53</sup>. In addition, the ATP-bound phosphorylated structure shows high fluctuations at the loop formed by the 2<sup>nd</sup> and 3<sup>rd</sup>  $\beta$ -strand, the  $\alpha$ DE loop, and the helix  $\alpha$ G and its adjoining loops. Compared with other Ser/Thr/Tyr kinase structures, the most variable regions in the IRAK4 structure are the  $\alpha$ DE loop, the activation segment, and the helix  $\alpha$ G and its adjoining loops<sup>32</sup>. These regions show a high degree of residue fluctuations in uIRAK4 and pIRAK4-ATP, indicating that these regions may be closely involved in substrate binding as well as in the initial inactivation mechanism of IRAK4 kinase. Notably, the high fluctuations observed in the loop regions of both the N and C lobes indicate the functional importance in IRAK4 kinase activity.

**ATP binding site.** To check whether phosphorylation alters the shape and size of the ATP binding pocket, we compared the IRAK4 structures of both the apo trajectories using the CASTp server and chimera tool. From this comparison, we observed that, in the case of phosphorylated structure, the ATP binding pocket was deeper and looked more accessible than in the unphosphorylated structure (Figure 6). The volumes of the ATP binding pockets in the phosphorylated and unphosphorylated structures were 1798 Å<sup>3</sup> (large pocket in pIRAK4) and 438.6 Å<sup>3</sup> (large pocket in uIRAK4) respectively. In addition, total solvent accessible surface areas (SASA) of the pIRAK4 and uIRAK4 structures were 13073.5 Å<sup>2</sup> and 12932.7 Å<sup>2</sup>, respectively. These results indicate that phosphorylation in the activation loop of IRAK4 KD alters the volume of the ATP binding pocket. Of the five  $\beta$ -strands in the N



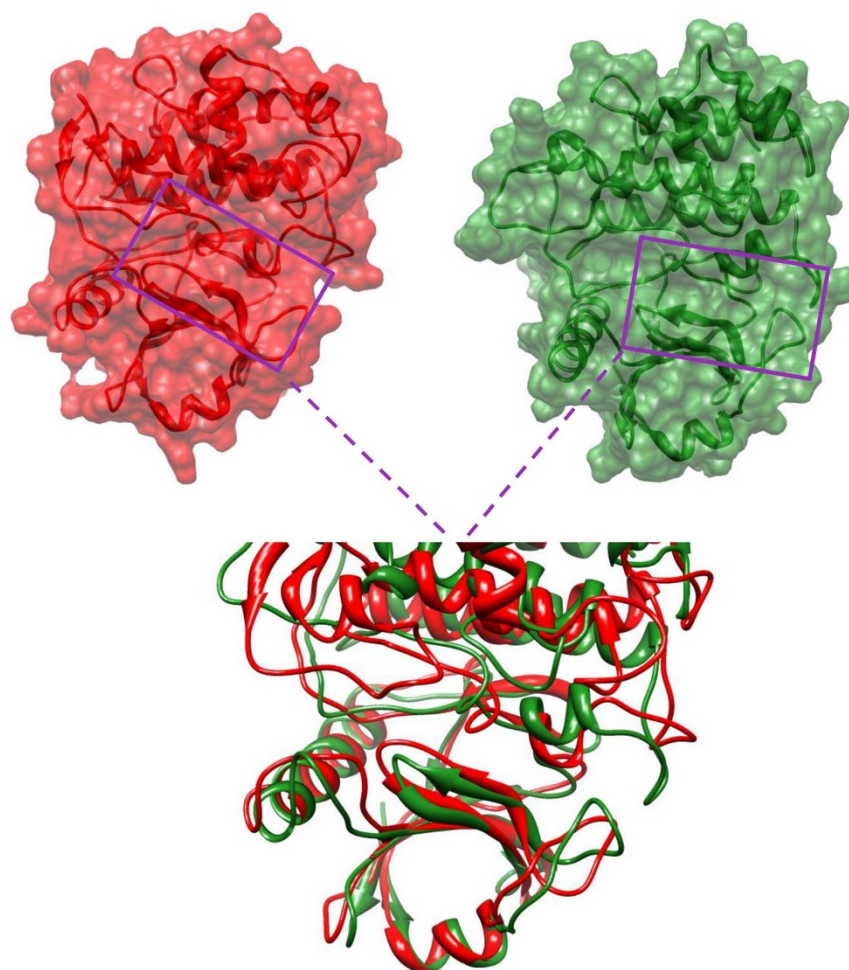
**Figure 4 | RMSD of Apo and ATP-bound complex during simulation.** RMSD of the backbone atoms with respect to initial structure of the apo and ATP-bound structures over 200-ns simulations.



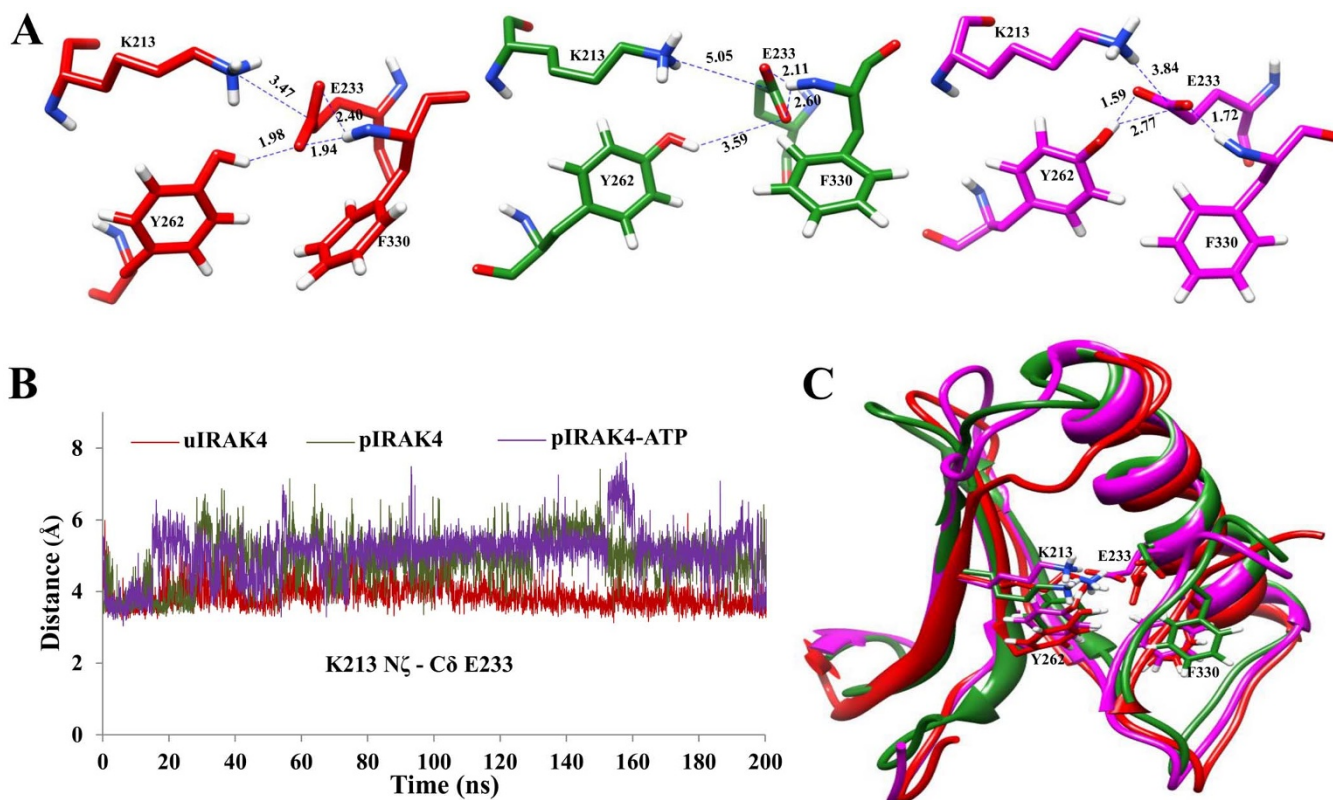
**Figure 5 | RMSF of Apo and ATP-bound complex during simulation.** RMSF plot of the backbone fluctuations obtained for apo and ATP-bound structures over last 100 ns (100 to 200 ns) simulations.

lobe, the first two  $\beta$ -strands shift slightly to the left; in particular, the G loop formed by the first two  $\beta$ -strands moves a little away to access the ATP molecule and thereby forms a deeper ATP binding pocket upon phosphorylation (Figure 6). This result suggests that the fully

phosphorylated IRAK4 KD undergoes conformational changes to enlarge the surface that is accessible for ATP binding. In addition, upon unphosphorylation (inactivation) of IRAK4 kinase, the N and C lobes move against each other via a hinge motion.



**Figure 6 | Structural difference in the ATP binding pocket.** Surface and ribbon representation (top) of unphosphorylated and phosphorylated apo structures with ATP binding pocket (magenta rectangle). Structural differences in the ATP binding pocket are shown by superimposing both the apo structures (bottom). Color codes are given in the Figure 2 legend.

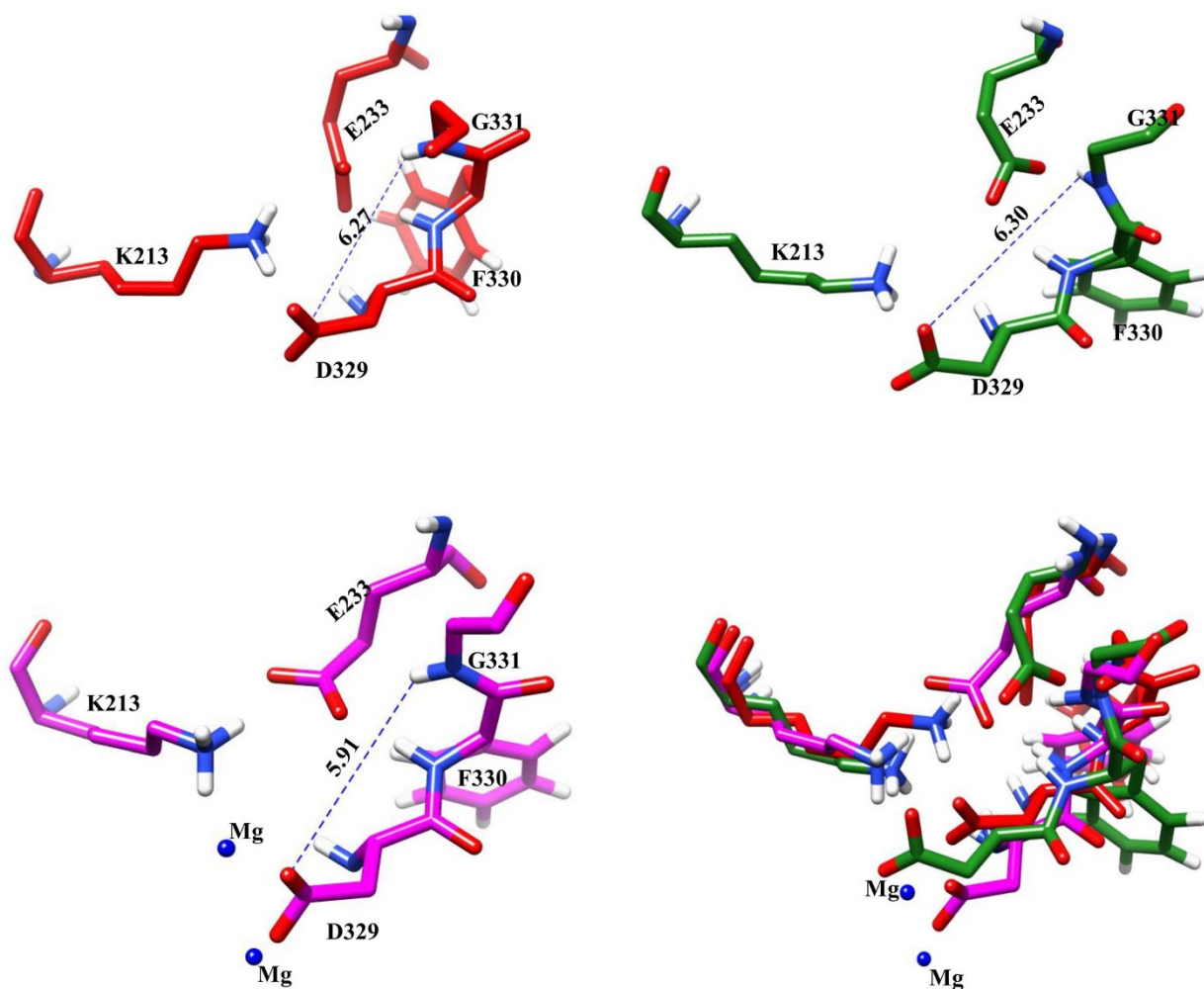


**Figure 7 | Helix  $\alpha$ C orientation.** (A) The structural representation of the key residues and their interactions are shown to understand the helix  $\alpha$ C orientation. Atoms are shown in stick representation. (B) The minimum distance is shown between the N $\zeta$  of K213 and C $\delta$  of E233. (C) Structural superimposition of the helix  $\alpha$ C from apo and ATP-bound structures. Color codes are given in the Figure 2 legend.

**Helix  $\alpha$ C orientation.** It has been consistently observed that helix  $\alpha$ C in the  $\beta$ -sheet-rich N lobe is an important mediator for the regulation mechanism of kinase, as this helix comprises a Glu residue that is crucial for the salt bridge interaction with the conserved Lys residue from the  $\beta$ 3 strand. This conserved Lys residue in turn coordinates with the  $\alpha$ - and  $\beta$ -phosphate groups of ATP<sup>54</sup>. Helix  $\alpha$ C often makes direct contact with the DFG motif, which is toward the N-terminal of the activation segment. Its role in kinase regulation is indicated by its proximity in the active state to the activation loop<sup>54</sup>. In both apo and ATP-bound IRAK4 structures, the helix  $\alpha$ C was found to have stable deviation from the initial structure during simulations (Supplementary Information Figure S4A). The IRAK4 active crystal structure shows two different conformations, which are opposite to those of the other Ser/Thr kinase, i.e., helix  $\alpha$ C-in contains no salt bridge, but helix  $\alpha$ C-out contains a salt bridge interaction between K213 from the  $\beta$ 3 strand and E233 from helix  $\alpha$ C<sup>34,39,40</sup>. Hence, the distance between the amino group of K213 and the carboxylate group of E233 was monitored throughout the simulations for both the apo and ATP-bound trajectories. Interestingly, the distance showed a remarkable result in terms of ATP binding, as well as the helix  $\alpha$ C. In the case of pIRAK4-ATP, the distance was 3.84 Å, whereas in both apo trajectories the distance was 5.05 Å for pIRAK4 and 3.47 Å for uIRAK4. Notably, in the pIRAK4-ATP structure, the distance was  $\sim$ 5.5 Å throughout the simulations up to 190 ns (Figures 7A and B). This result suggests that binding of ATP leads to further new interactions i.e., the negatively charged E233 coordinates with the positively charged Mg<sup>2+</sup> ion. In contrast, the positively charged K213 interacts with the negatively charged phosphates of ATP. In addition, we observed only helix  $\alpha$ C-in conformation in all three trajectories, since the distances mentioned above are too long to be considered salt bridges. This may be because the initial structure used for MD

simulation was in the helix  $\alpha$ C-in conformation. However, it is worth mentioning that the salt bridge distance between K213 and E233 is closer in uIRAK4 compared with pIRAK4.

Further, we were interested in the role of the gatekeeper residue Y262 in the phosphorylation of the kinase. Y262 hydrogen bonds with E233 in unphosphorylated IRAK4 and the ATP-bound phosphorylated IRAK4 model. In contrast, the distance was about 3.59 Å in phosphorylated IRAK4 model (Figure 7A). In pIRAK4, the residue distance was initially within 2.5 Å. However, after  $\sim$ 25 ns, the distance increased to  $\sim$ 4.5 Å (Supplementary Information Figure S4B). Therefore, the structural rearrangement occurs in the helix  $\alpha$ C with slight rotation and bending of the helix. However, this cannot be considered as the helix  $\alpha$ C-out conformation as we did not observe a K213-E233 salt bridge in the pIRAK4 structure. In addition, the RMSD of the superimposed helix  $\alpha$ C from both apo and ATP-bound models were 2.25 Å for uIRAK4 and pIRAK4, 2.30 Å for uIRAK4 and pIRAK4-ATP, and 2.04 Å for pIRAK4 and pIRAK4-ATP (Figure 7C). These RMSD values, in particular 2.25 Å for uIRAK4 and pIRAK4 and 2.04 Å for pIRAK4 and pIRAK4-ATP, suggest that both the helix  $\alpha$ C-in and helix  $\alpha$ C-out conformations are possible upon activation loop phosphorylation. In several kinases, the hinge movement of helix  $\alpha$ C away from the ATP binding site is required to switch from the catalytically active to the inactive state<sup>55</sup>. The above results suggest that unphosphorylated IRAK4 KD (inactive KD) exists in helix  $\alpha$ C-in conformation; once unphosphorylated IRAK4 becomes phosphorylated it may exist in both helix  $\alpha$ C-in and -out conformations (or more), and finally when ATP binds between N and C terminal lobes, may shift this equilibrium to the helix  $\alpha$ C-in position. In addition, E233 forms a hydrogen bond with the backbone amine of F330 in all three forms (Figure 7A). However, the residue distance between E233 and F330 in uIRAK4 is higher compared with pIRAK4 and pIRAK4-ATP (Supplementary Information



**Figure 8 | DFG motif conformation.** Structural representation of DFG motif from apo and ATP-bound structures are shown. Atoms are shown in stick representation. The distances between NH of G331 and oxygen of D329 are shown in blue dashes. The superimposition of DFG motif from apo and ATP bound structures are shown in the right bottom corner. Color codes are given in the Figure 2 legend.

Figure S4B). Hence, we speculate F330 may also be indirectly involved in the switching mechanism of helix  $\alpha$ C.

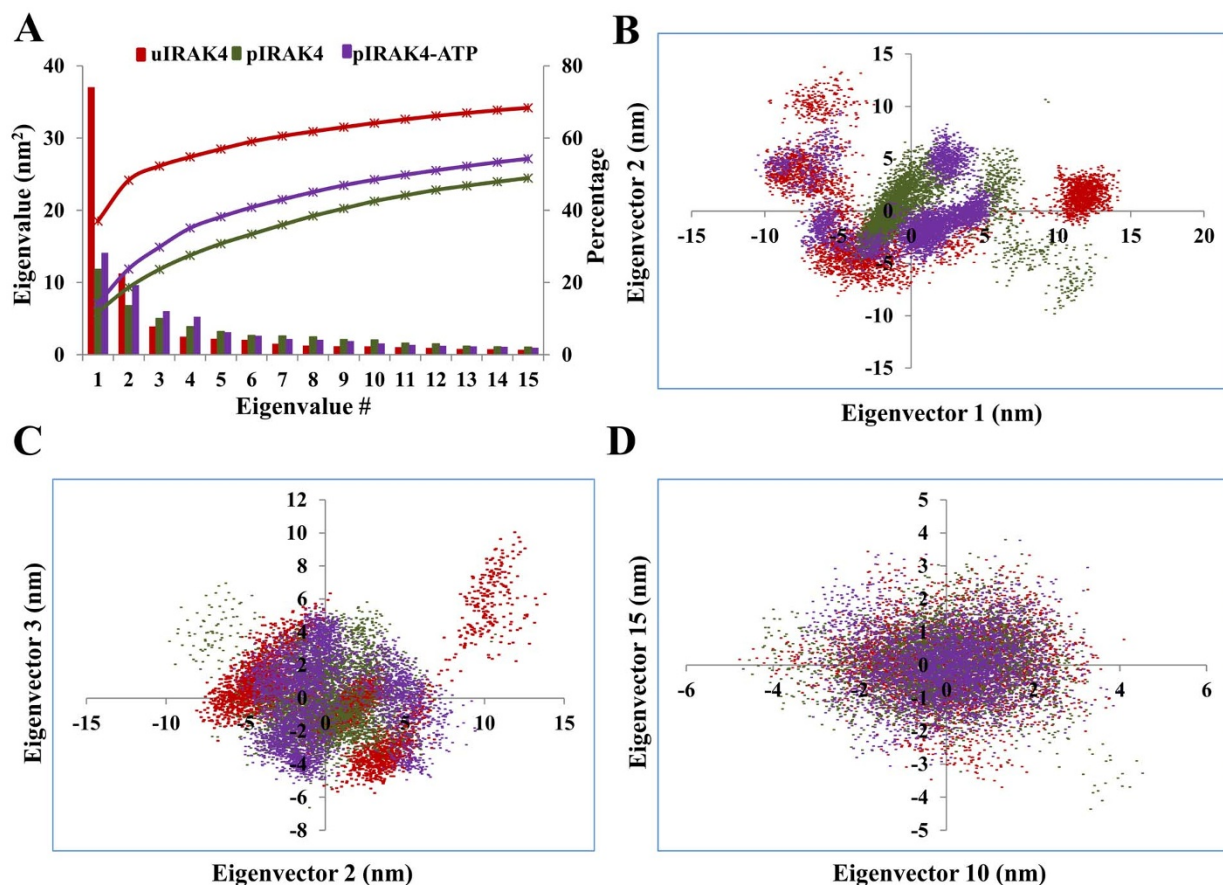
**DFG and HRD motifs.** Aspartic acids from DFG (329–331) and HRD (309–311) motifs are two critical residues responsible for the catalytic mechanism of many kinases. The structure-function relationship of the kinases, particularly the conformational changes in the DFG motif (i.e., the correlation of DFG flip with enzyme activity), has been widely studied. However, the exact mechanism of the DFG flip remains uncertain<sup>52,56,57</sup>. In active kinases, the DFG motif adopts DFG-in conformation as the Asp coordinates the magnesium ion, which in turn chelates ATP phosphates; this region of the kinase is called the magnesium-binding loop. The HRD motif acts as a base catalyst that deprotonates the Ser/Thr of the substrate side chain, and this region may be considered as a catalytic loop. Therefore, we were interested in understanding the DFG and HRD position in both apo and ATP-bound IRAK4 KD<sup>58</sup>.

In our study, we observed two different DFG conformations for both ATP bound and apo IRAK4 KD. In the case of the ATP-bound phosphorylated structure, D329 coordinated with the  $Mg^{2+}$  ion, thereby adopting DFG-in conformation, whereas in the apo structures D329 does not fully orient its side chain to coordinate with the  $Mg^{2+}$  ion (Figure 8). The side chain position of the phenylalanine in the ATP-bound form points toward the ATP binding site. In contrast, in both the apo structures uIRAK4 and pIRAK4, the side chain of

F330 points toward the helix  $\alpha$ C and helix  $\alpha$ E respectively. Therefore, we speculate that F330 may be involved in hydrophobic interactions of hydrophobic residues involving ATP binding. Kornev *et al*<sup>56</sup> suggested that the conformation of Asp in the active kinase relies on the hydrogen bond between the N-H group of Gly and the oxygen of Asp and that disruption of this hydrogen bond resulted in the possibility of DFG-flip. However, the Gly role in the DFG motif is not clearly defined. To test the above hypothesis for IRAK4 kinase, we measured the distance from the oxygen of D329 to the hydrogen of G331 from the three models: the distances in uIRAK4, pIRAK4, and pIRAK4-ATP were 6.27 Å, 6.30 Å, and 5.91 Å, respectively (Figure 8). Even though these distances cannot be considered as hydrogen bonding, these measurements may still be a reflection of the interaction, particularly upon ATP binding, to stabilize the position of aspartic acid<sup>45</sup>.

The other conserved motif in the kinase is the HRD motif (triad). In order to understand the active conformation of IRAK4, we investigated the HRD motif. In general, it is known that in Ser/Thr kinases, Arg from HRD interacts with the phosphate group from the activation loop<sup>35,36</sup>. In the case of the IRAK4 active crystal structure, R334 is the only residue that interacts with the phosphate group of T345 in the activation loop. However, during simulation, R310 maintained salt bridge interaction with pT345 in the ATP-bound phosphorylated IRAK4 structure. In contrast, this interaction was not observed in both the apo trajectories (Figure 3B). The minimum distances between T345 and R310 are shown in Figure 3A. In the case of





**Figure 9** | Projection of IRAK4 KD trajectories in phase space. (A) The first 15 eigenvectors of the covariance matrix and the lines with cross symbol represent the cumulative sum of the contribution to the total fluctuations. The clouds represent the projection of the trajectories of IRAK4 KD (B) eigenvector 1 and 2 (C), eigenvector 2 and 3 (D) eigenvector 10 and 15. We performed the analysis with 50-ps coordinates of the 200-ns trajectory; thus, we used 4000 frames of the trajectory for the projection. The color codes for the clouds are given in the Figure 2 legend.

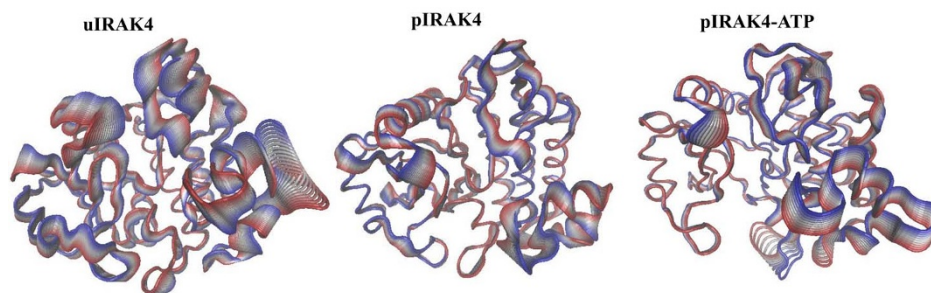
uIRAK4 and pIRAK4, the distance between R310 and T345 is too long. Additionally, we analyzed the position of invariant D311, in both apo and ATP-bound structures and the distance between D311 and nearby K313 was around 2–4 Å (data not shown). These distances suggest that D311 and K313 may form a salt bridge (Supplementary Information Figure S5). Additional interaction of D311 with T351 from the substrate-binding loop was observed in both apo structures (Supplementary Information Figure S5). D311 and K313 are greatly conserved in Ser/Thr kinases and the charged side chain groups of these residues are involved in the electrostatic stabilization of the transition state<sup>59</sup>. Our results strongly support the role of D311, which may trigger the hydroxyl group of the substrate, in the catalytic mechanism of IRAK4 kinase.

#### Collective modes of motion in apo and ATP-bound IRAK4 KD.

To better understand the conformational changes of the apo and ATP-bound IRAK4 KD structures, the MD trajectories of each system were inspected with the principal components. Approximately, the first 15 modes contribute greatly to the collective motions, as only the first 15 eigenvectors showed values greater than 1 nm<sup>2</sup>. The first 15 collective modes for each system with cumulative percentage are shown in Figure 9A. The application of PCA components for uIRAK4, pIRAK4, and pIRAK4-ATP structures and the motions of simulations showed 68%, 49%, and 54%, respectively, and accounted for the first 15 eigenvectors (Figure 9A). From these results, we observed that the internal motion of both apo and ATP-bound IRAK4 KD is limited to subspace with fewer dimensions. The projection of IRAK4 KD, defined by eigenvectors 1 and 2, 2 and 3, 10 and 15 from each MD

trajectory, allows visualization onto the essential space (Figures 9B, C, and D). In particular, the projection of MD trajectories onto eigenvectors 1 and 2 showed that all the three trajectories were sufficient to reach distinct minima with small energy barriers. Distribution of all three systems showed that both apo and ATP-bound IRAK4 undergo conformational changes. In particular, uIRAK4 projection, which indicates the mechanism of inactivation of IRAK4 kinase, undergoes large conformational changes. The pIRAK4 model shows dissimilar eigenvectors compared to the uIRAK4 and pIRAK4-ATP trajectories. However, all the three trajectories showed extensive sampling of phase space until the system reached a different equilibrium on the energy landscape. Figures 9C and D show that the motions of the protein attain their equilibrium fluctuations within 15 eigenvectors. Thus, the three well-defined clusters seen in Figures 9B, 9C, and 9D indicate the conformational changes that may occur upon phosphorylation and ATP binding to the unphosphorylated IRAK4 kinase as well as upon inactivation of IRAK4 kinase.

The subdomain motions within IRAK4 were analyzed by the corresponding motion mode of the first eigenvector, as seen in Figure 10. The animation of motions for 200-ns MD trajectories of the three systems can be seen in the movies S1–S3 for eigenvector 1 and S4–S6 for eigenvector 2. As expected, large domain motions were observed for the apo trajectories compared with those for the ATP-bound form. In unphosphorylated IRAK4 kinase, overall, the N and C lobes move toward each other around the hinge region (Figure 10, Supplementary Information Movie S1). The hinge region connects the N and C lobes and interacts with the edge of the ATP molecule. Therefore, the N and C lobes independently contribute to the surface



**Figure 10 | Eigenvector fluctuations from PCA analysis.** The principal motion projected along first eigenvector1 is superimposed sequentially with 20 frames of the whole trajectory. Red and blue colors represent from minimum to maximum, respectively. The protein is represented in the tube model.

of the ATP binding site, in particular the activation loop in the activation segment moves toward the C-lobe helices and helix  $\alpha$ G moves toward the activation segment. The second dominant motion of uIRAK4 is in the opposite direction to the C lobe, including the activation segment with the N lobe (Supplementary Information Movie S4). The N lobe and helix  $\alpha$ G, and its adjoining loops, show more pronounced dynamics in unphosphorylated IRAK4 KD, suggesting that these regions may play a role in the inactivation dynamics of IRAK4 kinase.

In the phosphorylated IRAK4 trajectory, the first dominant mode can be described as the breathing mode, in which helix  $\alpha$ C and the G loop from the N lobe move against the C lobe, particularly against the activation segment and helix  $\alpha$ G and its adjoining loop (Figure 10, Supplementary Information Movie S2). Strikingly, the activation segment showed less pronounced motion, which is consistent with the RMSF calculations. Similar motions were found in other kinases<sup>53,60</sup>. The most significant motion of the second dominant mode in pIRAK4 was the N lobe moving against helix  $\alpha$ G and its adjoining loops and activation segment (Supplementary Information Movie S5). The first dominant mode in ATP-bound IRAK4 involved helix  $\alpha$ G and its adjoining loops, and the  $\alpha$ DE loop from the C lobe moved against the activation segment, in particular substrate binding site, while the N lobe remained stable (Figure 10, Supplementary Information Movie S3). The motion of the second dominant mode in pIRAK4-ATP involved helix  $\alpha$ G and its adjoining loop, and the  $\alpha$ DE loop moved in the opposite direction to the activation segment (Supplementary Information Movie S6). The dominant motions of helix  $\alpha$ G and the  $\alpha$ DE loop upon ATP binding to the phosphorylated IRAK4 suggested the importance of helix  $\alpha$ G and the  $\alpha$ DE loop during the phosphorylation transition state catalyzed by the IRAK4 kinase.

## Discussion

Protein kinases have been characterized as one of the essential components in the cell signaling pathways of all eukaryotic organisms<sup>61,62</sup>. Several studies have indicated that IRAK4 may be a suitable therapeutic drug target for inhibiting the production of proinflammatory genes via the TLR/IL-1R signaling pathways<sup>6,21,28,63</sup>. In this study, comparative molecular dynamics simulations were performed with apo and ATP-bound IRAK4 KD structures. At the 200-ns timescale of IRAK4 KDs, the activation segment of the unphosphorylated IRAK4 (uIRAK4) structure showed a high degree of diversity compared to that of the phosphorylated IRAK4 kinase (pIRAK4 and pIRAK4-ATP) (Figure 2A). In the uIRAK4 model, the activation loop appears to be located in close proximity to the C lobe, which may allow closing of the substrate docking site, indicating the inactive nature of the kinase. However, in active IRAK4 kinase, the Arg triad may pull the whole activation loop, thereby exposing the p + 1 loop for the IRAK1 substrate. A wide variety of inactive conformations of the activation segment have previously been observed, and in some cases the activation segment appeared to be quite disordered<sup>48</sup>. Crystal coordinates have shown that active IRAK4 KD activation

loop phosphorylation is similar to that of other Ser/Thr/Tyr kinases<sup>32</sup> such as ERK2 (MAP kinase), cAPK, CDK2, LCK, and IRK. Although IRAK4 kinase is a Ser/Thr kinase, the phosphate interaction pattern of active IRAK4 kinase is closely related to that of Tyr kinases such as LCK, IRK, IGF1RK, and JAK3, i.e., single primary phosphate ligand interaction with Arg or Lys, equivalent to Arg334 of IRAK4 kinase<sup>32</sup>. Interestingly, the pIRAK4-ATP model showed direct interaction of pT345 with R310 and R334 throughout the simulations, indicating that, upon ATP binding to IRAK4 kinase, R310 and R334 are mutually complementary (Figure 3A). Most Ser/Thr kinases have been observed to possess two or three phosphate-binding residues, and in most RD kinases, the arginine from the HRD (YRD in PKA) motif interacts with the primary phosphorylation site upon activation<sup>36,48,62</sup>. However, in pIRAK4, pT345 showed direct interaction with R334 alone, whereas in uIRAK4 T345/R334 was quite distant, indicating that the lobe separation in the inactivation mechanism of IRAK4 kinase is facilitated through disruption of the interaction of pT345 with R310 and R334 (Figure 3A). Upon ATP binding, hydrogen bond interaction was observed between the phosphate group of pS346 and Q228 (polar residue) from helix  $\alpha$ C, whereas in apo IRAK4 these residues are relatively distant (Figure 3B) (Supplementary Information Figure S3). Mutation of this residue reduces IRAK4 kinase activity up to 50%<sup>33</sup>. This type of interaction has previously been observed in most of the Ser/Thr kinases such as active MAPK kinase (R68-pT183), in which the PKB (H196-pT309) basic residue from helix  $\alpha$ C interacts with the phosphorylated residue from the activation loop, whereas Tyr kinases lack such interactions<sup>52</sup>. This analysis reveals that, upon activation of IRAK4 kinase, apart from the Arg triad, the Q228 from helix  $\alpha$ C may also be involved in switching the helix  $\alpha$ C back to close proximity to the DFG motif at the N-terminus of the activation segment, thereby facilitating the proper orientation for the DFG flip.

The ATP binding site is located between two lobes of the kinase domain. ATP binding volumes varied upon phosphorylation and unphosphorylation of IRAK4 KD, suggesting that upon inactivation of IRAK4 kinase, the motion of the hinge region may allow closure of the N and C lobes. In PKA, the relative orientation of the N and C lobes varies from open in the inactive conformation to closed in the active conformation via hinge motion, enabling tight binding of ATP molecule<sup>61</sup>. The helix  $\alpha$ C in the N lobe of the kinase is a unique and dynamic regulatory element. In active Ser/Thr kinase, an electrostatic interaction was observed between Lys from the  $\beta$ 3 strand and Glu from helix  $\alpha$ C, indicating the helix  $\alpha$ C-in conformation, whereas the inactive kinase maintains the helix  $\alpha$ C-out conformation (i.e., no salt bridge interaction is formed between Lys and Glu). The opposite is observed in IRAK4 kinase; in both the apo and ATP-bound models, the helix  $\alpha$ C-in conformation was maintained, as no salt bridge interaction was formed between the conserved K213 and E233 residues (throughout in the simulations the residue distance remained  $\sim$ 3.5–5.5 Å) (Figure 7). The superimposed RMSD differences in the helix  $\alpha$ C suggest that it may appear in two different conformations in solution upon phosphorylation of the IRAK4



kinase (helix  $\alpha$ C-in and -out). In Src-family tyrosine kinases, the intramolecular domains SH3 and SH2 are involved in the regulatory mechanism<sup>64,65</sup>, and it may be worth noting that the helix  $\alpha$ C switch in IRAK4 kinase may be triggered by the death domain as well as the undetermined domain (via intramolecular interactions) of the IRAK4 kinase, or by an adaptor protein such as MyD88 (via intermolecular interactions). The “gatekeeper” residue plays an important role in the activation of the kinase, and although 77% of human kinases have large gatekeeper residues (Leu, Met, and Phe), 21% are tyrosine kinases, with small residues (Thr, Val). IRAK4 has a Tyr residue as a gatekeeper; comparison with the 400 kinases has shown that this Tyr is unique to the IRAK family<sup>32</sup>, as well as the plant receptor-like kinases (RLKs)<sup>66</sup>. Recently, it has been reported that kinases belonging to the IRAK family are very similar in structure and function to plant RLKs such as LYK3<sup>67</sup>. In IRAK4 KD, the hydrogen bond was maintained between Y262 and E233 in both the inactive and active models (Figure 7A), suggesting that the back pocket of the ATP-binding site may be completely inaccessible to ATP-competitive inhibitors in both active and inactive IRAK4 kinases. However, mutation of the gatekeeper residue may have substantial physiological significance. Mutation of the gatekeeper residue has been shown to be associated with drug resistance in chronic myeloid leukemia and several solid tumor patients<sup>68</sup>. Because the back pocket of the ATP-binding site is blocked in IRAK4 kinase, the extended front pocket of the ATP-binding area may have a direct impact on inhibitor design, compared to other protein kinases<sup>32</sup>. This may be true for all members of the IRAK family, as all have Tyr as the gatekeeper residue.

IRAK4 kinase showed a DFG-in conformation when ATP was present and a DFG-out conformation for apo IRAK4 (Figure 8). The spatial arrangement of the HRD catalytic motif revealed the importance of R310 and D311 upon ATP binding and confirmed the active state of the enzyme (Supplementary Information Figure S5). The major structural differences in active IRAK4 KD are the loop between helices  $\alpha$ D and  $\alpha$ E (276–282) and helix  $\alpha$ G and its adjoining loops (383–394 and 407–423). The position of helix  $\alpha$ G is at a considerable distance from the active site and helix  $\alpha$ EF, with an approximately 10-Å shift compared to other Ser/Thr/Tyr kinases<sup>6</sup>. In addition, although the N-terminal extension of IRAK4 associates with the core structure of the kinase, no sequence similarity has been observed between the N-terminal extensions of various kinases<sup>32,69</sup>. Using PCA analysis, we observed the collective motions of three forms of IRAK4 KD. In the pIRAK4-ATP structure, pronounced dynamics were observed in the C lobe and in the loop regions of the N lobe. More strikingly, the dominant motions for the  $\alpha$ DE loop and helix  $\alpha$ G with its adjoining loops were observed upon activation (pIRAK4-ATP) and inactivation (uIRAK4). In particular, the dominant motion of helix  $\alpha$ G in active (pIRAK4-ATP) IRAK4 kinase shows that helix  $\alpha$ G and its adjoining loop play a crucial role in adapting the substrate for the catalytic mechanism upon activation (Figure 10). The dominant fluctuations observed for helix  $\alpha$ G in the uIRAK4 model also suggest that helix  $\alpha$ G plays a crucial role in the inactivation mechanism of IRAK4 kinase and that the loop regions within the IRAK4 kinase play a major role in the inter dynamics of IRAK4 kinase.

We strongly believe that this type of computational approach is necessary to understand the structure-function relationship in protein kinases. To our knowledge, this is the first study exploring the dynamic nature of both apo and ATP-bound IRAK4 kinase, which may provide some new insights for further research and drug development.

## Methods

**Apo and ATP-bound IRAK4 KD preparation.** Crystal coordinates of active IRAK4 KD in complex with inhibitor benzimidazole were downloaded from the Protein Data Bank (PDB) (PDB entry 2NRU). There is a crystal structure available for active IRAK4 KD in complex with  $\beta$ , $\gamma$ -imidoadenosine 5'-triphosphate lithium salt hydrate

(AMP-PNP) (PDB entry 2OID). However, this structure has several missing residues, and the superimposition of both 2NRU and 2OID showed a root mean square deviation (RMSD) of 0.5 Å. Hence, we used the 2NRU structure for further studies. First, we removed all the ligand and water molecules from the 2NRU crystal structure. In the 2NRU structure, only the residues T345 and S346 at the activation loop are phosphorylated; to construct fully phosphorylated IRAK4 KD, we additionally phosphorylated T342 in the activation loop using *Discovery studio visualizer* (Discovery Studio 2.1, Accelrys Inc., San Diego, CA, USA), as these three autophosphorylation sites are responsible for IRAK4 kinase activity<sup>33</sup>. To build the complex with ATP (pIRAK4-ATP), we superimposed the fully phosphorylated structure on to the crystal structure of 2OID, retrieved the AMP-PNP molecule from 2OID, and modified the AMP-PNP ligand as the ATP molecule. Additionally, we included two magnesium ions ( $Mg^{2+}$ ) based on the crystal structure of phosphorylase kinase (PDB entry 1PHK). We have included two magnesium ions because IRAK4 kinase needs more than one  $Mg^{2+}$  ion for its catalytic mechanism<sup>70,71</sup>. To construct pIRAK4, we removed the ATP and  $Mg^{2+}$  ions from the complex structure pIRAK4-ATP. Although the phosphorylation mechanism of the kinase has substantial physiological importance<sup>72</sup>, no crystal coordinates are currently available for inactive IRAK4 KD, and it is not yet computationally possible to model the conformation of the inactive kinase with physiological significance. Hence, we constructed an unphosphorylated IRAK4 (uIRAK4) KD to represent inactive IRAK4 KD. To construct uIRAK4, we replaced the phosphothreonine and phosphoserine residues at T342, T345, and S346 with Thr and Ser from pIRAK4. A similar approach has been reported previously<sup>45,53</sup>. Eventually, we subjected the three different systems (uIRAK4, pIRAK4, and pIRAK4-ATP) to further dynamics simulations.

**Molecular dynamics (MD) simulations.** Classical MD simulation was performed for uIRAK4, pIRAK4, and pIRAK4-ATP using the Gromacs 4.6<sup>73,74</sup> package. The Gromos43a1 force field and PRODRG2.5<sup>75</sup> server were used to generate the topologies for protein and ATP, respectively. The uIRAK4, pIRAK4, and pIRAK4-ATP systems were solvated in a corresponding box under the periodic boundary conditions using a distance of 1.0 nm from the protein to the surface of the box. To neutralize the system, counter ions were added accordingly. Each system was subjected to energy minimization using the steepest descent integrator without constraints for 1000 steps. Following the minimization step, systems were equilibrated under NVT (canonical ensemble) and NPT (isothermal-isobaric ensemble) conditions for 100 ps at 300 K after applying position restraints to the protein. Finally, a 200-ns production run was performed under NPT conditions by removing position restraints. The temperature and pressure of the system were maintained using the Berendsen weak-coupling method. Lennard-Jones potentials were used for Vander walls interactions, and electrostatic interactions were handled by PME calculations with a cut-off for the real space term of 0.9 nm<sup>76</sup>. The LINCS algorithm was used to constrain all the bonds<sup>77</sup>. A 2-fs time step was applied and 2-ps final coordinates were saved. Most of the analyses were performed using Gromacs inbuilt tools with 50-ps coordinate. The final low energy conformation structures were extracted from last 10 ns (190–200 ns) of each simulation, and this structure has been considered as a representative structure whenever necessary to compare the three different systems. The RMSD and root mean square fluctuation (RMSF) calculations were done by least square fit. All visualizations were performed using Chimera<sup>78</sup>, VMD<sup>79</sup>, and PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC). All the graphs were plotted using Microsoft Excel.

**Principal component analysis.** To obtain dominant and collective modes of the protein from the overall dynamics of the MD trajectory, PCA was performed. PCA is based on the construction of a mass-weighted covariance matrix of the atom displacement, and this covariance matrix is diagonalized to extract a set of eigenvectors and eigenvalues that reflect concerted motion of the molecule<sup>80–82</sup>. Eigenvectors represent the direction of motion, whereas the corresponding eigenvalues represent the amplitudes in those directions. The IRAK4 coordinates were extracted from each time frame of 50-ps MD trajectories from each system. Overall rotational and translational movements in the MD trajectory were removed by a least-squares fit superimposition on the average structure<sup>83</sup>. The Gromacs inbuilt tool `g_covar` was used to yield the eigenvalues and eigenvectors by calculating and diagonalizing the covariance matrix, whereas the `g_anaeig` tool was used to analyze and plot the eigenvectors.

1. Akira, S. & Takeda, K. Toll-like receptor signalling. *Nat. Rev. Immunol.* **4**, 499–511 (2004).
2. Akira, S. TLR signaling. *Curr. Top. Microbiol. Immunol.* **311**, 1–16 (2006).
3. O'Neill, L. A., Bryant, C. E. & Doyle, S. L. Therapeutic targeting of Toll-like receptors for infectious and inflammatory diseases and cancer. *Pharmacol. Rev.* **61**, 177–197 (2009).
4. Makkouk, A. & Abdelnoor, A. M. The potential use of Toll-like receptor (TLR) agonists and antagonists as prophylactic and/or therapeutic agents. *Immunopharmacol. Immunotoxicol.* **31**, 331–338 (2009).
5. Gosu, V., Basith, S., Kwon, O. P. & Choi, S. Therapeutic applications of nucleic acids and their analogues in Toll-like receptor signaling. *Molecules* **17**, 13503–13529 (2012).
6. Wang, Z., Wesche, H., Stevens, T., Walker, N. & Yeh, W. C. IRAK-4 inhibitors for inflammation. *Curr. Top. Med. Chem.* **9**, 724–737 (2009).



7. Janssens, S. & Beyaert, R. Functional diversity and regulation of different interleukin-1 receptor-associated kinase (IRAK) family members. *Mol. Cell* **11**, 293–302 (2003).
8. Gosu, V., Basith, S., Durai, P. & Choi, S. Molecular evolution and structural features of IRAK family members. *PLoS One* **7**, e49771 (2012).
9. Medzhitov, R. *et al.* MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. *Mol. Cell* **2**, 253–258 (1998).
10. Fitzgerald, K. A. *et al.* Mal (MyD88-adaptor-like) is required for Toll-like receptor-4 signal transduction. *Nature* **413**, 78–83 (2001).
11. Martin, M. U. & Wesche, H. Summary and comparison of the signaling mechanisms of the Toll/interleukin-1 receptor family. *Biochim. Biophys. Acta* **1592**, 265–280 (2002).
12. Kobayashi, K. *et al.* IRAK-M Is a Negative Regulator of Toll-like Receptor Signaling. *Cell* **110**, 191–202 (2002).
13. Brikos, C., Wait, R., Begum, S., O'Neill, L. A. J. & Saklatvala, J. Mass Spectrometric Analysis of the Endogenous Type I Interleukin-1 (IL-1) Receptor Signaling Complex Formed after IL-1 Binding Identifies IL-1RACp, MyD88, and IRAK-4 as the Stable Components. *Mol. Cell. Proteomics* **6**, 1551–1559 (2007).
14. Li, S., Strelow, A., Fontana, E. J. & Wesche, H. IRAK-4: a novel member of the IRAK family with the properties of an IRAK-kinase. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5567–5572 (2002).
15. Suzuki, N. *et al.* Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature* **416**, 750–756 (2002).
16. Picard, C. *et al.* Pyogenic Bacterial Infections in Humans with IRAK-4 Deficiency. *Science* **299**, 2076–2079 (2003).
17. Medvedev, A. E. *et al.* Cutting edge: expression of IL-1 receptor-associated kinase-4 (IRAK-4) proteins with mutations identified in a patient with recurrent bacterial infections alters normal IRAK-4 interaction with components of the IL-1 receptor complex. *J. Immunol.* **174**, 6587–6591 (2005).
18. Ku, C. L. *et al.* Selective predisposition to bacterial infections in IRAK-4-deficient children: IRAK-4-dependent TLRs are otherwise redundant in protective immunity. *J. Exp. Med.* **204**, 2407–2422 (2007).
19. Ku, C. L. *et al.* Inherited disorders of human Toll-like receptor signaling: immunological implications. *Immunol. Rev.* **203**, 10–20 (2005).
20. Lye, E., Mirtsos, C., Suzuki, N., Suzuki, S. & Yeh, W. C. The role of interleukin 1 receptor-associated kinase-4 (IRAK-4) kinase activity in IRAK-4-mediated signaling. *J. Biol. Chem.* **279**, 40653–40658 (2004).
21. Koziczak-Holbro, M. *et al.* IRAK-4 kinase activity is required for interleukin-1 (IL-1) receptor- and toll-like receptor 7-mediated signaling and gene expression. *J. Biol. Chem.* **282**, 13552–13560 (2007).
22. Koziczak-Holbro, M. *et al.* Solving the IRAK-4 enigma: application of kinase-dead knock-in mice. *Ernst Schering Found. Symp. Proc.* **3**, 63–82 (2007).
23. Kim, T. W. *et al.* A critical role for IRAK4 kinase activity in Toll-like receptor-mediated innate immunity. *J. Exp. Med.* **204**, 1025–1036 (2007).
24. Lye, E., Dhanji, S., Calzascia, T., Elford, A. R. & Ohashi, P. S. IRAK-4 kinase activity is required for IRAK-4-dependent innate and adaptive immune responses. *Eur. J. Immunol.* **38**, 870–876 (2008).
25. Yang, K. *et al.* Human TLR-7-, -8-, and -9-mediated induction of IFN- $\alpha$ / $\beta$  and - $\lambda$  Is IRAK-4 dependent and redundant for protective immunity to viruses. *Immunity* **23**, 465–478 (2005).
26. Isnardi, I. *et al.* IRAK-4- and MyD88-dependent pathways are essential for the removal of developing autoreactive B cells in humans. *Immunity* **29**, 746–757 (2008).
27. Barrat, F. J., Meeker, T., Chan, J. H., Guiducci, C. & Coffman, R. L. Treatment of lupus-prone mice with a dual inhibitor of TLR7 and TLR9 leads to reduction of autoantibody production and amelioration of disease symptoms. *Eur. J. Immunol.* **37**, 3582–3586 (2007).
28. Sun, M., Yang, P., Du, L., Yang, Y. & Ye, J. The role of interleukin-1 receptor-associated kinases in Vogt-Koyanagi-Harada disease. *PLoS One* **9**, e93214 (2014).
29. Cao, Z., Henzel, W. J. & Gao, X. IRAK: a kinase associated with the interleukin-1 receptor. *Science* **271**, 1128–1131 (1996).
30. Muzio, M. IRAK (Pelle) Family Member IRAK-2 and MyD88 as Proximal Mediators of IL-1 Signaling. *Science* **278**, 1612–1615 (1997).
31. Wesche, H. *et al.* IRAK-M is a novel member of the Pelle/interleukin-1 receptor-associated kinase (IRAK) family. *J. Biol. Chem.* **274**, 19403–19410 (1999).
32. Wang, Z. *et al.* Crystal Structures of IRAK-4 Kinase in Complex with Inhibitors: A Serine/Threonine Kinase with Tyrosine as a Gatekeeper. *Structure* **14**, 1835–1844 (2006).
33. Cheng, H. *et al.* Regulation of IRAK-4 kinase activity via autophosphorylation within its activation loop. *Biochem. Biophys. Res. Commun.* **352**, 609–616 (2007).
34. Kuglstatler, A. *et al.* Cutting Edge: IL-1 receptor-associated kinase 4 structures reveal novel features and multiple conformations. *J. Immunol.* **178**, 2641–2645 (2007).
35. Johnson, L. N. & Lewis, R. J. Structural basis for control by phosphorylation. *Chem. Rev.* **101**, 2209–2242 (2001).
36. Nolen, B., Taylor, S. & Ghosh, G. Regulation of protein kinases; controlling activity through activation segment conformation. *Mol. Cell* **15**, 661–675 (2004).
37. De Bondt, H. L. *et al.* Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595–602 (1993).
38. Goldberg, J., Nairn, A. C. & Kuriyan, J. Structural basis for the autoinhibition of calcium/calmodulin-dependent protein kinase I. *Cell* **84**, 875–887 (1996).
39. Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602–609 (1997).
40. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595–602 (1997).
41. Gan, W., Yang, S. & Roux, B. Atomistic view of the conformational activation of Src kinase using the string method with swarms-of-trajectories. *Biophys. J.* **97**, L8–L10 (2009).
42. Berteotti, A. *et al.* Protein conformational transitions: the closure mechanism of a kinase explored by atomistic simulations. *J. Am. Chem. Soc.* **131**, 244–250 (2009).
43. Groban, E. S., Narayanan, A. & Jacobson, M. P. Conformational changes in protein loops and helices induced by post-translational phosphorylation. *PLoS Comput. Biol.* **2**, e32 (2006).
44. Banavali, N. K. & Roux, B. Anatomy of a structural pathway for activation of the catalytic domain of Src kinase Hck. *Proteins* **67**, 1096–1112 (2007).
45. Gruszczynski, P., Obuchowski, M. & Kazmierkiewicz, R. Phosphorylation and ATP-binding induced conformational changes in the PrkC, Ser/Thr kinase from *B. subtilis*. *J. Comput. Aided. Mol. Des.* **24**, 733–747 (2010).
46. Shan, Y., Arkhipov, A., Kim, E. T., Pan, A. C. & Shaw, D. E. Transitions to catalytically inactive conformations in EGFR kinase. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7270–7275 (2013).
47. Steichen, J. M. *et al.* Global consequences of activation loop phosphorylation on protein kinase A. *J. Biol. Chem.* **285**, 3825–3832 (2010).
48. Johnson, L. N., Noble, M. E. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149–158 (1996).
49. Hubbard, S. R., Mohammadi, M. & Schlessinger, J. Autoregulatory mechanisms in protein-tyrosine kinases. *J. Biol. Chem.* **273**, 11987–11990 (1998).
50. Knighton, D. R. *et al.* Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407–414 (1991).
51. Russo, A. A., Jeffrey, P. D. & Pavletich, N. P. Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat. Struct. Biol.* **3**, 696–700 (1996).
52. Krupa, A., Preethi, G. & Srinivasan, N. Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. *J. Mol. Biol.* **339**, 1025–1039 (2004).
53. Cheng, S. & Niv, M. Y. Molecular dynamics simulations and elastic network analysis of protein kinase B (Akt/PKB) inactivation. *J. Chem. Inf. Model.* **50**, 1602–1610 (2010).
54. Jeffrey, P. D. *et al.* Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**, 313–320 (1995).
55. Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **109**, 275–282 (2002).
56. Kornev, A. P., Haste, N. M., Taylor, S. S. & Eyck, L. F. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17783–17788 (2006).
57. Adams, J. A. Kinetic and catalytic mechanisms of protein kinases. *Chem. Rev.* **101**, 2271–2290 (2001).
58. Gibbs, C. S. & Zoller, M. J. Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions. *J. Biol. Chem.* **266**, 8923–8931 (1991).
59. Madhusudan Akamine, P., Xuong, N. H. & Taylor, S. S. Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nat. Struct. Biol.* **9**, 273–277 (2002).
60. Cheng, Y., Zhang, Y. & McCammon, J. A. How does activation loop phosphorylation modulate catalytic activity in the cAMP-dependent protein kinase: a theoretical study. *Protein Sci.* **15**, 672–683 (2006).
61. Jura, N. *et al.* Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Mol. Cell* **42**, 9–22 (2011).
62. Taylor, S. S. & Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **36**, 65–77 (2011).
63. Picard, C., Casanova, J. L. & Puel, A. Infectious diseases in patients with IRAK-4, MyD88, NEMO, or IkappaBalpha deficiency. *Clin. Microbiol. Rev.* **24**, 490–497 (2011).
64. Lerner, E. C. *et al.* Activation of the Src family kinase Hck without SH3-linker release. *J. Biol. Chem.* **280**, 40832–40837 (2005).
65. Moarefi, I. *et al.* Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* **385**, 650–653 (1997).
66. Samaddar, S. *et al.* Autophosphorylation of gatekeeper tyrosine by symbiosis receptor kinase. *FEBS Lett.* **587**, 2972–2979 (2013).
67. Klaus-Heisen, D. *et al.* Structure-function similarities between a plant receptor-like kinase and the human interleukin-1 receptor-associated kinase-4. *J. Biol. Chem.* **286**, 11202–11210 (2011).
68. Tamborini, E. *et al.* A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient. *Gastroenterology* **127**, 294–299 (2004).
69. Lei, M., Robinson, M. A. & Harrison, S. C. The active conformation of the PAK1 kinase domain. *Structure* **13**, 769–778 (2005).
70. Hekmat-Nejad, M., Cai, T. & Swinney, D. C. Steady-state kinetic characterization of kinase activity and requirements for Mg<sup>2+</sup> of interleukin-1 receptor-associated kinase-4. *Biochemistry* **49**, 1495–1506 (2010).
71. Yu, L. *et al.* Role of Mg<sup>2+</sup> ions in protein kinase phosphorylation: insights from molecular dynamics simulations of ATP-kinase complexes. *Molecular Simulation* **37**, 1143–1150 (2011).



72. Keshwani, M. M. *et al.* Cotranslational cis-phosphorylation of the COOH-terminal tail is a key priming step in the maturation of cAMP-dependent protein kinase. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E1221–1229 (2012).
73. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).
74. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
75. Schüttelkopf, A. W. & van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).
76. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N [center-dot] log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
77. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
78. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
79. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–38 (1996).
80. Amadei, A., Linssen, A. B. & Berendsen, H. J. Essential dynamics of proteins. *Proteins* **17**, 412–425 (1993).
81. Yamaguchi, H., van Aalten, D. M., Pinak, M., Furukawa, A. & Osman, R. Essential dynamics of DNA containing a cis.syn cyclobutane thymine dimer lesion. *Nucleic Acids Res.* **26**, 1939–1946 (1998).
82. van Aalten, D. M., Findlay, J. B., Amadei, A. & Berendsen, H. J. Essential dynamics of the cellular retinol-binding protein—evidence for ligand-induced conformational changes. *Protein Eng.* **8**, 1129–1135 (1995).
83. Amadei, A., Linssen, A. B., de Groot, B. L., van Aalten, D. M. & Berendsen, H. J. An efficient method for sampling the essential subspace of proteins. *J. Biomol. Struct. Dyn.* **13**, 615–625 (1996).

## Acknowledgments

This work was supported by the Mid-Career Researcher Program through the National Research Foundation of Korea, funded by the Ministry of Education, Science and Technology (2012R1A2A2A02016803) and also partly supported by a grant from the Priority Research Centers Program (NRF 2012-0006687).

## Author contributions

V.G. designed and performed the experiments. V.G. and S.C. wrote the main manuscript text. All authors reviewed the manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Gosu, V. & Choi, S. Structural dynamic analysis of apo and ATP-bound IRAK4 kinase. *Sci. Rep.* **4**, 5748; DOI:10.1038/srep05748 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>