

Community analysis of large-scale molecular dynamics simulations elucidated dynamics-driven allostery in Tyrosine kinase 2

Metaxia Vlassi¹ and Nastazia Lesgidou¹

¹Ethniko Kentro Ereunas Physikou Epistemon Demokritos

July 18, 2023

Abstract

TYK2 is a non-receptor tyrosine kinase, member of the Janus kinases (JAK), with a central role in several diseases, including cancer. The JAKs' catalytic domains (KD) are highly conserved, yet the isolated TYK2-KD exhibits unique specificities. In a previous work, using molecular dynamics (MD) simulations of a catalytically-impaired TYK2-KD variant (P1104A) we found that this amino-acid change of its JAK-characteristic insert (α FG), acts at the dynamics level. Given that structural dynamics is key to allosteric activation of protein kinases, in this study we applied a long-scale MD simulation and investigated an active TYK2-KD form in the presence of adenosine 5'-triphosphate and one magnesium ion that represents a dynamic and crucial step of the catalytic cycle, in other protein kinases. Community analysis of the MD trajectory shed light, for the first time, on the dynamic profile and dynamics-driven allosteric communications within the TYK2-KD during activation and revealed that α FG and amino-acids P1104, P1105 and I1112 in particular, hold a pivotal role and act synergistically with a dynamically coupled communication network of amino-acids serving intra-KD signaling for allosteric regulation of TYK2 activity. Corroborating our findings, most of the identified amino-acids are associated with cancer-related missense/splice-site mutations of the *Tyk2* gene. We propose that the conformational dynamics at this step of the catalytic cycle, coordinated by α FG, underlies TYK2-unique substrate recognition and accounts for its distinct specificity. In total, this work adds to knowledge towards an in-depth understanding of TYK2 activation and may be valuable towards a rational design of allosteric TYK2-specific inhibitors.

Community analysis of large-scale molecular dynamics simulations elucidated dynamics-driven allostery in Tyrosine kinase 2

Running title: Dynamics-driven allostery in TYK2 kinase

Nastazia Lesgidou and Metaxia Vlassi *

Institute of Biosciences & Applications, National Center for Scientific Research "Demokritos", Athens, 15310, Greece

* Corresponding author: Metaxia Vlassi (meta@bio.demokritos.gr)

Acknowledgments:

We acknowledge support of this work by the projects: "INSPIRED-The National Research Infrastructures on Integrated Structural Biology, Drug Screening Efforts and Drug target functional characterization" (MIS 5002550) implemented under the Action "Reinforcement of the Research and Innovation Infrastructure" and "SANITURA-Target Identification and Development of Novel Approaches for Health and Environmental Applications" (MIS 5002514) implemented under the Action for the Strategic Development on the Research and Technological Sectors, both funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund). This work was also supported by computational time granted from the National Infrastructures for Research and Technology S.A. (GRNET) in the National HPC facility - ARIS - under project

IDs: KIN_IMMUNMD-II and KIN_IMMUNMD-III. N.L. acknowledges support through the Operational Programme «Human Resources Development, Education and Lifelong Learning» (NSRF 2014-2020) in the context of the Act “Enhancing Human Resources Research Potential by undertaking a Doctoral Research” Sub-action 2: IKY Scholarship Programme for PhD candidates in the Greek Universities by the program “Human Resource Development, Education and Lifelong Learning”.

Abstract

TYK2 is a non-receptor tyrosine kinase, member of the Janus kinases (JAK), with a central role in several diseases, including cancer. The JAKs’ catalytic domains (KD) are highly conserved, yet the isolated TYK2-KD exhibits unique specificities. In a previous work, using molecular dynamics (MD) simulations of a catalytically-impaired TYK2-KD variant (P1104A) we found that this amino-acid change of its JAK-characteristic insert (α FG), acts at the dynamics level. Given that structural dynamics is key to allosteric activation of protein kinases, in this study we applied a long-scale MD simulation and investigated an active TYK2-KD form in the presence of adenosine 5'-triphosphate and one magnesium ion that represents a dynamic and crucial step of the catalytic cycle, in other protein kinases. Community analysis of the MD trajectory shed light, for the first time, on the dynamic profile and dynamics-driven allosteric communications within the TYK2-KD during activation and revealed that α FG and amino-acids P1104, P1105 and I1112 in particular, hold a pivotal role and act synergistically with a dynamically coupled communication network of amino-acids serving intra-KD signaling for allosteric regulation of TYK2 activity. Corroborating our findings, most of the identified amino-acids are associated with cancer-related missense/splice-site mutations of the *Tyk2* gene. We propose that the conformational dynamics at this step of the catalytic cycle, coordinated by α FG, underlies TYK2-unique substrate recognition and accounts for its distinct specificity. In total, this work adds to knowledge towards an in-depth understanding of TYK2 activation and may be valuable towards a rational design of allosteric TYK2-specific inhibitors.

Keywords :

Tyrosine kinase 2; molecular dynamics simulations; protein kinases; allostery; community map analysis; Janus kinases; cancer mutations

1. Introduction

Eukaryotic protein kinases (EPKs) are highly dynamic, ATP (adenosine 5'-triphosphate)-dependent enzymes that catalyze phosphorylation of their downstream protein-substrates, thus acting as “switches” of signaling pathways with pivotal role in regulating key processes in human cells. Through their catalytic cycle they have to transition from the inactive to active state and vice-versa and their conformational plasticity and concerted motions within their catalytic domains are essential for regulating kinase activity¹⁻². The catalytic domains of EPKs share high sequence and fold similarities³⁻⁴ therefore identification of residues that determine distinct specificities is challenging.

Tyrosine kinase 2 (TYK2) is a non-receptor tyrosine kinase that belongs to the four-membered, Janus family of kinases (JAKs) (the others being JAK1-3) and in an interplay with JAK1 and JAK2, is involved in cell signaling in response to various immunoregulatory cytokines as part of the JAK-STAT signaling pathway depending on the cytokine and receptor complex⁵⁻⁶. TYK2 plays a central role in both innate and adaptive immune and inflammatory responses as well as in tumor immune-surveillance and its aberrant function is associated with the pathogenesis of several diseases including (auto)immune and inflammatory diseases as well as cancer; actually, *Tyk2* has emerged as an oncogene⁷⁻¹³. It has also been reported that the TYK2-STAT3 signaling pathway, in addition to cancer, is implicated in Alzheimer’s disease¹⁴, whereas tyrosine phosphorylation at a different site uniquely by the isolated catalytic domain of TYK2, has been reported to negatively control STAT3 activity¹⁵. Unraveling the atomic details of TYK2 kinase activity and the determinants of its specificity, is therefore, of major importance.

Structurally, TYK2 follows a JAK-characteristic architecture; namely, it is divided into four domains: the FERM and SH2 domains that are responsible for receptor binding¹⁶, a pseudo-kinase domain (pKD) that

regulates its kinase activity¹⁷⁻¹⁸ and the catalytic domain (KD).

The TYK2-KD is responsible for its tyrosine kinase activity and adopts the EPK characteristic bilobate fold (see e.g., in¹⁹ and Fig. S1). In brief, it is structurally divided into two lobes connected with a linker (hinge region); namely, a small N-terminal lobe consisting of a 5-stranded β -domain and the regulatory C-helix that is key to the ON-OFF switch of kinase activity, and a large mainly α -helical C-terminal lobe (Fig. S1). The ATP molecule and protein-substrates must bind in the cleft between the two lobes. ATP in particular, must bind deep in the cleft in such a way that its adenine-ring is buried in a conserved, transiently formed dynamic hydrophobic pocket². The N-lobe provides most of the residues that are involved in correct, catalytically competent positioning of ATP, including a conserved lysine located at the β 3 strand (K930 in TYK2), which serves to coordinate the ATP α - and β -phosphates and is stabilized by a salt-bridge with a conserved glutamate residue of the regulatory C-helix (KE salt-bridge; K930-E947 in TYK2) in active kinases as well as a glycine-rich loop (G-loop) that contributes to aligning the ATP-phosphates towards the substrate⁴. The C-lobe on the other hand, encompasses elements that are characteristic of EPKs³; namely, the activation loop and a co-evolved helical subdomain that comprises three α -helices (α G- α H- α I; GHI-subdomain). Tyrosine phosphorylation of its activation loop (AL) at Y1054-Y1055, is required for TYK2 activation²⁰. The C-lobe also provides several regions and motifs both conserved and specific that are involved in substrate-binding and catalysis, such as the HRD- and DFG-motifs³, the APE-motif that connects AL with the GHI-subdomain²¹ and the substrate-binding, P+1 loop (Fig. S1). The ¹⁰⁶⁴PVFWYA¹⁰⁶⁹ region of the TYK2 P+1 loop in particular, is characteristic of tyrosine-specific kinases⁴ (PTKs). Both the N- and C-lobes provide residues serving non-covalent communication between the lobes, through the formation of transiently assembled spatial motifs in the dynamic inner hydrophobic core of active EPKs; namely the regulatory R-spine (assembled upon AL phosphorylation) and the catalytic C-spine (transiently assembled upon ATP-binding) as well as additional bridging residues that link the spines and combined, serve dynamics-driven allosteric kinase activation^{2,22-23}.

The TYK2-KD C-lobe has an addition of an α -helix near the substrate-binding site, the FG-helix (Fig. S1), which is characteristic of JAKs and plays a central role in their (auto)activation²⁴. However, the atomic details of this, remain elusive. One amino-acid change of the TYK2 α FG, in particular; namely, a Pro-1104 to Ala substitution, results in a catalytically impaired TYK2 protein (P1104A) and is associated with increased cancer risk albeit protective against autoimmune diseases.^{13, 25-30}

In a previous work, and as a first step towards the elucidation of the atomic-details of altered TYK2 functionality, we applied all-atom molecular dynamics (MD) simulations and studied the dynamics of the functionally impaired P1104A-KD in comparison with that of the wild type, in their apo-forms³¹. MD simulations offer a very powerful alternative to experimental techniques and have been widely used to study dynamic processes, including allostery in protein kinases.³²⁻³⁵ Analysis of the MD trajectories indicated that this particular α FG amino-acid change quenches the dynamics of the resulting TYK2-KD and promotes closed, collapsed and inaccessible to ATP conformations³¹. Given that structural dynamics underlies dynamics-driven allosteric activation of protein kinases^{2,23} and that movements within the TYK2-KD are restricted by its other subdomains in the context of the full-length protein in its autoinhibited state¹⁷⁻¹⁸, these results prompted us to investigate the dynamics of the TYK2-KD during the activation process, that remain unexplored until today.

To this end, in the present study, we applied a microsecond-scale MD simulation (3 μ s) and investigated the dynamics of the AL-phosphorylated (at Y1054-Y1055) wild type TYK2-KD in the presence of ATP and one magnesium ion (ATP.1Mg²⁺) that represents the activation process and corresponds to a crucial and highly dynamic pre phosphoryl-transfer step of the catalytic cycle, in other protein kinases.^{32, 35-36}

Analysis of the MD-trajectory, including community network analysis, shed light, for the first time, on the dynamic profile and dynamics-driven allosteric communications within the TYK2-KD in response to AL phosphorylation and ATP.1Mg²⁺-binding, and revealed that the FG-helix and amino-acids P1104, P1105 and I1112 in particular, hold a pivotal role and are part of a dynamically coupled allosteric communication network of amino-acids across the lobes, serving dynamic bridging of the spines and intra-KD allosteric

signaling for dynamics-driven allosteric regulation of TYK2 activity. The MD results suggest that this is most-probably facilitated by the dynamic nature of several areas of the TYK2-KD, including the β 3- α C and α H- α I connecting regions, in a TYK2-unique manner. Corroborating the functional significance of the identified allosteric network of side-chains, most of the key residues (more than 40%) identified in this study, are associated with cancer-related *Tyk2* missense or splice-site mutations, as revealed by screening of related databases. We propose that this pre phosphoryl-transfer step of the catalytic cycle is crucial for TYK2 activation and that the conformational dynamics at this step, coordinated by α FG, is essential for substrate recognition and thus, accounts for its distinct specificity (e.g. for TYK2-specific sites of STAT3). In total, this work adds to knowledge towards an in-depth understanding of TYK2 activation and provides information that may be valuable towards a structure-based design of allosteric TYK2-specific inhibitors.

2. Methods

Molecular Dynamics simulations

All-atom molecular dynamics (MD) simulations were performed using the GROMACS software package (v. 5.1.4).³⁷⁻³⁸ Due to lack of known crystal structures of the simulated TYK2-KD form, 3D-modeling of the initial conformation was necessary, prior to the MD simulation. The known crystal structure of an ADP.2Mg²⁺-bound form of a TYK2-KD (rcsb PDB:4gvj³⁹; and Fig. S1) was used as template, for this purpose. Due to amino-acid changes in 4gvj, *in-silico* mutagenesis using Swiss-PdbViewer⁴⁰ was applied to obtain the wild-type TYK2-KD form. The coordinates for the ADP molecule and of one magnesium ion were kept, whereas missing coordinates for the ATP γ -phosphate group and for the phosphate groups of the phosphorylated AL tyrosine residues, Y1054p and Y1055p, were transferred from the PDB:1atp and PDB:6dbk⁴¹ crystal structures, respectively.

The simulations were carried out in explicit water using periodic dodecahedron boxes of TIP3P-water molecules⁴² extending 10 Å from protein atoms, to solvate the protein system. Periodic boundaries were applied to minimize edge effects. The solvated system was subsequently neutralized with counter-ions, optimized by steepest descent energy minimization and equilibrated by restrained MD simulations for 1 ns, in two steps: namely, the protein/ATP.1Mg²⁺ atoms were harmonically restrained to allow the solvent to equilibrate for 500 ps in the NVT ensemble at constant temperature, followed by 500 ps in the NPT ensemble at a pressure of 1 bar. The optimization phase was followed by the 3 μ s-long unrestrained MD simulation in the NPT ensemble (productive run). All simulations were performed at a constant temperature of 300 K with separate coupling of protein and non-protein atoms. In order to better prepare the system, a much shorter (\sim 40 ns) preparatory MD run (and its corresponding equilibration steps) was carried out at 150K, prior to the 300K simulation. The AMBER99SB-ILDN⁴³ forcefield as implemented in GROMACS was used in the simulations, as we have previously shown that it produces more reliable MD results⁴⁴. The AMBER parameters for phosphotyrosines and for ATP with a net charge of -4e, were adapted from the Bryce R, AMBER Parameter Database (<http://pharmacy.man.ac.uk/amber>). The PME⁴⁵ method for the treatment of the long-range electrostatic interactions, a cut-off radius of 10 Å for all non-bonded interactions and a 2 fs time step for integration of the potential function, were used in all simulations.

The 3 μ s-long MD simulation was carried out in the National HPC facility -ARIS-, supported by a computational time grant from the National Infrastructures for Research and Technology S.A. (GRNET) (under project IDs: KIN_IMMUNMD-II and -III).

Analysis of the MD-trajectory

GROMACS³⁸ was also used for various preliminary analyses of the MD-trajectory based on several metrics, such as RMSD and RMSF calculations, analysis of geometric parameters (e.g. angles, dihedrals and distances) and conformation clustering (as in³¹). In brief, convergence of the MD-trajectory was assessed by monitoring the root-mean-square deviation (RMSD) of several atoms from the initial positions, whereas conformational flexibility was evaluated by means of root-mean-square fluctuation (RMSF) calculations. The RMSF values are a measure of atomic fluctuations; i.e. atomic displacements relative to their mean positions along the MD-trajectory. Conformation clustering was applied to obtain representative MD-snapshots at

various time-frames of the simulation.

Additional analyses of the MD-trajectory related to identifying concerted atomic motions included:

Principal Component (PC) analysis

PC analysis was applied to eliminate the noise and to identify the dominant modes of atomic movements during the MD simulation (as in³¹). The PC-analysis was performed in the range of 2-3 μ s of the MD-trajectory and the movements of C α atoms and side-chains (a representative atom) were treated separately. Components PC1-PC100 obtained from the independent PC-analyses were saved as new trajectories (PCA-trajectories) and were used in the subsequent analyses of the MD trajectory.

Cross-correlation analysis

Cross-correlation analysis of the MD trajectory was used in order to examine the degree of correlation of the atomic motions within the simulated TYK2-KD. This type of analysis is based on the calculation of cross correlation matrices the elements of which, namely, the cross-correlation coefficients (Cij) between each pair of atoms i,j (see eq. $\text{Corr}_{i,j}$ in³¹), take values ranging from -1 to +1 and can identify atomic movements in a collective manner (semi-rigid bodies) either in the same or in opposite directions (positive or negative values, respectively). Cij values close to ± 1 indicate highly correlated atomic displacements, whereas values close to 0 are indicative of uncorrelated motions. The Bio3D software⁴⁶⁻⁴⁷ and the PCA-trajectories obtained as described above, were used for this analysis.

Community map analysis; node-Betweenness analysis

Community map analysis was performed also using the Bio3D software⁴⁶⁻⁴⁷ and the separate, C α and side-chain, PCA-trajectories. Contact maps produced using an atom-atom distance cut-off of [?] 10Å for at least 75% of the analyzed simulation time and the Girvan-Newman algorithm⁴⁸ as implemented in Bio3D, were used for this purpose. The Girvan-Newman method is a graph-based network approach that is based on the edge-betweenness centrality measure (details in⁴⁸). In brief, the edge-betweenness centrality of an individual residue is defined as the number of the shortest paths connecting other residue pairs that pass through it along the MD trajectory, thus providing an estimate of the influence of this residue on communication (modularity). Communities of residues are characterized by high modularity values; i.e. residues in the same community share dense connection, whereas residues of different communities have sparse or no connections at all. The results are then plotted on a 2D network-graph with nodes representing communities and edges representing the strength of allosteric coupling between communities. Two independent community analyses, based on only positive (Cij>0.5) or both positive and negative cross-correlation (|Cij|>0.5) matrix calculations, were carried out for each one of the C α and side-chain PCA-trajectories.

Model illustrations were made using the PyMOL Molecular Graphics System⁴⁹.

Multiple sequence Alignment

ClustalW⁵⁰ was used for multiple alignment of JAK-KD amino-acid sequences, retrieved from the Uniprot-database⁵¹.

Cancer mutations search

Cancer-associated mutations were obtained by screening the cBioPortal for Cancer Genomics⁵².

3. Results and Discussion

As already mentioned, our previous MD work on the apo form of the TYK2-KD in comparison with its functionally impaired P1104A variant, indicated that this α FG amino-acid alteration has long range effects and acts at the dynamics rather than at the structural level of the TYK2-KD³¹. It has long been recognized that dynamics and concerted motions within the KD domains of EPKs during their catalytic cycle (e.g. in response to ATP-binding) play a pivotal role in their allosteric activation.^{2,23} In addition, AL phosphorylation

is necessary for kinase activation and has been reported to also result in increased KD dynamics in other kinases.⁵³

Since the dynamics and allosteric communications within the TYK2-KD in response to ATP-binding or AL phosphorylation remain unexplored until today, in this study, we applied a much longer, microsecond-scale (3 μ s) all-atom MD simulation and studied the TYK2-KD in the presence of ATP and one magnesium ion (ATP.1Mg²⁺). Magnesium ions are essential for kinase activity; namely for transferring the ATP γ -phosphate (phosphoryl-transfer) to protein-substrates. Binding of ATP and one Mg²⁺ ion in particular, has been reported to represent a crucial and highly dynamic step of the catalytic cycle of other protein kinases^{32,35-36}. In addition, the AL-phosphorylated form (Y1054p-Y1055p) was used in order to simulate a fully activated TYK2-KD and to investigate the effect of AL phosphorylation on its dynamics.

Due to lack of crystal structures of this TYK2-KD form, a 3D-model of the starting conformation was produced prior to the MD simulation, based on a known crystal structure of its ADP.2Mg²⁺ bound form (Fig. S1), as described in *Methods*.

A first evaluation of the MD trajectory was carried out by monitoring the root-mean-square deviation (RMSD; see *Methods*) of the protein-C α , ATP-adenine and magnesium atoms, relative to their initial positions along the 3 μ s MD-trajectory (Fig. S2). As reflected by relatively small RMSD values (\sim 2 Å), a convergence of the MD-trajectory albeit dynamic as expected for an activated protein kinase, was reached after 2 μ s of the simulation (Fig. S2).

3.1 Proper ATP positioning in the inner hydrophobic core but transiently formed K930(β 3)-E947(α C) salt-bridge, were observed in the simulated TYK2-KD

To investigate the details of the ATP-binding site, several related geometric parameters were monitored along the MD-trajectory. As already mentioned, kinase activity requires correct positioning of the ATP molecule deep into the ATP-binding groove, via a hydrogen bond of the ATP-adenine with a conserved protein anchoring site; namely, the backbone oxygen of a conserved glutamate residue at the hinge region (E979 in TYK2³⁻⁴). Monitoring of this distance (d in Fig. S3) showed that the ATP molecule retained its proper anchoring to the protein (in hydrogen-bonding distance; $d \sim$ 2.5 Å) throughout the 3 μ s of the MD-trajectory (Fig. S3B). On the other hand, a “breathing” of the entrance of the ATP pocket was observed, as reflected by a two peak distribution of its angle, ϑ (Fig. S3), indicative of a switching between ‘open’ and ‘closed’ conformations of the G-loop relative to the catalytic loop. As already mentioned, relative motions of the N-lobe with respect to the C-lobe are essential for the catalytic activity of protein kinases^{2, 54} and their restriction, e.g. by pKD in the monomeric autoinhibited state, has been proposed to block TYK2 activation¹⁷⁻¹⁸. Indeed, a rigid ATP-binding pocket and a collapsed G-loop, were suggested by our previous MD work on the catalytically impaired P1104A variant³¹.

Accordingly, the side-chain of the conserved ATP-binding β 3 lysine residue, K930, oscillated between conformations either in close proximity to the ATP α/β -phosphates or in hydrogen-bonding distances to its α C-glutamate partner in forming the K930-E947 salt-bridge along the MD-trajectory (Fig. 1). Temporary disruption of the KE salt-bridge has also been observed in the monomeric form of AL-phosphorylated dimeric receptor-PTKs; e.g. in the case of EGFR⁵³. The conformational swapping of the K930 side-chain observed here, was more clear in the range of 2-3 μ s of the simulation (Fig. 1) and this time-range was used in subsequent analyses of the MD-trajectory.

3.2. Disruption of the C-lobe, APE-R salt-bridge and increased dynamics of the phosphorylated activation loop and of the α H- α I connecting region, were detected in the simulated TYK2-KD

Another hallmark of active EPKs has been proposed to be a conserved salt-bridge in the C-lobe between the APE-glutamate (E1071) and a conserved arginine residue (R1159) located at the loop connecting the α H and α I helices, that mediates communication of the activation loop with the C-lobe GHI-subdomain with an important role in the catalytic activity of other protein kinases.^{21,55}

Interestingly, monitoring of the E1071-R1159 (APE-R) distance along the MD-trajectory, revealed that

this salt-bridge is mainly disrupted in the TYK2-KD simulated here (Fig. 1), reflecting a highly dynamic character of the TYK2 AL and α H- α I loops in response to AL phosphorylation and/or ATP.1Mg²⁺-binding.

To further investigate the details of conformational flexibility within the simulated TYK2-KD, RMSF calculations were carried out (see *Methods*). As shown in Fig. 2, increased backbone dynamics (as reflected by higher RMSF values) were indeed detected for AL and the α H- α I connecting fragment as well as for several other, mainly N-lobe regions. More specifically, almost all the N-lobe loops including the β 3- α C, β 2- β 3 and β 4- β 5 connecting fragments (and the aforementioned G-loop), exhibited large backbone fluctuations in the simulated TYK2-KD (Fig. 2). In addition to AL and α H- α I regions, and albeit the C-lobe overall appears to be more rigid compared to the N-lobe, a third C-lobe area emerged from the RMSF analysis as also highly dynamic; namely, the region connecting the α FG and α G helices (Fig. 2), which is related to substrate-binding in JAKs (based on substrate-mimicking binding of suppressors of cytokine signaling proteins; SOCS⁵⁶). This finding is in line with the notion that increased flexibility of substrate-binding regions is characteristic of PTKs⁵⁷.

Taken together our MD results so far, reveal that ATP.1Mg²⁺-binding and AL phosphorylation increase the TYK2-KD dynamics, in line with the established notion that elevated dynamics and concerted side-chain movements are essential to coordinating activation of protein kinases.^{2,53}

3.3 Cross-correlation analysis of the MD-trajectory revealed a lack of correlated atomic movements between the activation loop and the GHI-subdomain

To identify intra-KD correlated movements in response to ATP.1Mg²⁺-binding and AL phosphorylation, a cross-correlation analysis of the MD trajectory was first applied (see *Methods*). Side-chain atoms were taken into account in this analysis as separate entities, due to their key role in mediating allosteric communications, necessary for kinase activity². An illustration of highly positively-correlated atomic motions in the time-frame of 2-3 μ s of the simulation, focused on the aforementioned dynamic C-lobe regions, is shown in Fig. 3.

As shown in Fig. 3, there is a disruption of communication between the GHI-subdomain (α H- α I loop) and AL (APE region), as reflected by the absence of correlation (lack of connecting lines) between the APE-R side-chains, E1071 and R1159 (see also §3.2) as well as of a related network of hydrophobic interactions expected in the region²¹ involving the side-chains of A1156, R1159, A1081 and W1085 (Figs. 3 and S4). This is most probably due to the replacement of a central to stabilizing these interactions, hydrophobic side-chain²¹ (Leu277^{PKA}) by an alanine in TYK2 (A1156). Indeed, the predicted APE-R shielding side-chain contact, A1156-A1081, is disrupted and replaced by an unforeseen interaction of residues A1156 and R1159, in the simulated TYK2-KD (Figs. 3 and S4). Interestingly, Ala-1156 is TYK2-unique, whereas position 1156 preserves its hydrophobic character in the other JAK members (i.e. it is either a Pro or Val; Fig. S5), suggesting that the disruption of the interaction network at the APE-R region observed here, is TYK2-specific and crucial for its specific activity. Corroborating this idea, valine substitutions for A1156 or A1081, predicted by our analysis to restore the hydrophobic interaction network in the region, are detected in several adenocarcinomas.⁵²

Taken together our MD results so far, suggest a unique dynamic profile for the TYK2-KD in response to ATP.1Mg²⁺-binding and AL phosphorylation.

3.4. Community map analysis of the MD-trajectory revealed the dynamic profile of the TYK2-KD in response to ATP.1Mg²⁺-binding and AL phosphorylation

To investigate this issue, we carried out a community map analysis of the MD-trajectory. This type of analysis combined with microsecond-scale MD simulations has been widely used to analyze allosteric communications and to obtain the dynamic profiles of various proteins including protein kinases, under various conditions.^{32, 34-35} One of these studies; namely, a community analysis work of MD trajectories of ATP/Mg²⁺-bound PKA forms, divided the catalytic domain of this Ser/Thr-specific kinase into dynamic communities (i.e. transient clusters of residues that move as semi-rigid bodies), each with identifiable func-

tion, and suggested that this dynamic architecture is a common feature of EPKs.³² Variations of dynamic profiles depending, for example, on the functional state of the kinase or characteristic to each kinase family, are expected. Community analysis of the MD-trajectory of this study, is therefore expected to shed light on allosteric communications within the TYK2-KD in response to ATP.Mg²⁺-binding and AL phosphorylation and to unravel the details of its characteristic dynamic profile during activation.

To this end, cross-correlation matrices were combined with community network analysis based on the Girvan-Newman algorithm as implemented in Bio3D⁴⁶⁻⁴⁷ and as described in *Methods*. The Girvan-Newman⁴⁸ method offers an effective way for describing allosteric interactions as it provides a quantitative estimate of allosteric coupling based on the edge-betweenness measure from graph-theory (see *Methods*). Main-chain and side-chain atomic motions as well as positive only and positive/negative (All) cross-correlations, were treated separately. Hereafter, we will mainly focus on positively-correlated atomic motions unless otherwise noted, as they provide information of long-range collective movements in the same direction (see also Fig. 3). Communities were defined by the C α -based community analysis (MC), whereas the sidechain-based (SC) community analysis provided a more detailed information of allosteric communications; namely, it provided information on community bridging residues (a residue is defined as “bridging” if its C α and sidechain atoms belong to different communities) that play a crucial role in dynamics-driven allosteric intradomain communications.^{2, 32} Nomenclature and coloring of the identified communities was according to their structurally and/or functionally equivalent PKA communities³², whereas extra communities were named based on TYK2 specific secondary structure elements they are built around.

3.4.1. Dynamic profile of the simulated TYK2-KD

The community network map analysis of the MD-trajectory, detected 14 highly interconnected communities of dynamic amino-acids with positively-correlated atomic motions within the simulated TYK2-KD, which are depicted using three types of illustration in Fig. 4A (a bar plot of community bringing residues is shown, in Fig. S6). The N-lobe is partitioned in five communities (Com-N, -A, -A1, -A2 and -B), whereas the remaining nine communities are formed by C-lobe regions, including the regulatory C-helix (Fig. 4A). Among the identified communities, eight are structurally equivalent to major EPK communities (Com-A to -H), whereas six extra TYK2 communities emerged from our analysis that seem to be TYK2/JAK specific (Com-A1, -A2, -N, -FG, -I, and -P+1). The identified TYK2-KD communities are described in detail, below.

N-lobe Communities:

Com-A (in red) is centered around the five-stranded β -sheet (Fig. 4Aa-b) and is the main N-lobe and the third largest MC-community identified in the TYK2-KD simulated here (Fig. S7). It comprises residues that are involved in correct positioning of ATP, such as hydrophobic residues lining the upper surface of the ATP-adenine-binding pocket (e.g. C-spine residues, V911 and A928; Fig. 4B), the conserved ATP-binding residues, E979 and K930 as well as the entire G-loop. It thus, corresponds to the ATP-binding community that is involved in the dynamic assembly of the N-lobe β -structure and in correct positioning of the ATP-adenine ring, which is a prerequisite for committing the kinase to catalysis. Com-A forming residues also include the R-spine residue, Y962 and hydrophobic residues L976 and M978 (Fig. 4B/ “*R spine Shell*”), which correspond to amino-acids of the inner hydrophobic core that dynamically bridge the spines of activated protein kinases.^{2,22} The side-chain of the “gatekeeper” methionine residue, M978, in particular, blocks access to a hydrophobic pocket adjacent to the ATP-binding site and is mainly a threonine in other PTKs, with an important role in regulating their auto-activation⁵⁸. Interestingly, our community analysis identified M978 as a “bridging” residue with the regulatory community, Com-C (described below), strongly suggesting a similar role of its side-chain in the regulation of TYK2 (and JAKs) kinase activity. In support of this idea, residue M978 together with Y962 have been reported to act as a regulatory switch in TYK2¹⁹ and are conserved in JAKs (Fig. S5).

Com-B (in orange) is a small community that is built around the flexible β_3 - α C linker including the N-terminal portion of the C-helix (Fig. 4Aa-b). It serves dynamic allosteric coupling between the ATP-adenosine-binding community, Com-A and Com-F that corresponds to the activation community (two edges

of the orange sphere in Fig. 4Ac), suggesting an important role of the Com-B region in coupling ATP-binding with TYK2 activation. Corroborating this idea, the TYK2 β_3 - α C region is a cancer hotspot. For example, Com-B residues of dynamic nature, such as G943 and its adjacent S942, are unique to TYK2 among the JAKs (Fig. S5) and *Tyk2* point mutations affecting their dynamics and thus, the dynamics of the C-helix; namely, S942L and G943C, have been identified in mixed and lung cancers, respectively⁵². Combined these observations, suggest a key role for the dynamics of the β_3 - α C loop in α C-IN positioning and TYK2 activation. Notably, an important role of disorder-to-order transitions of this region in activation, has been reported for the monomeric form of other dimeric PTKs; eg. the receptor tyrosine kinase, EGFR⁵³.

Com-A1 (in salmon) is a TYK2/JAK-specific community and encompasses the β_2 - β_3 bulge, situated on top of the N-lobe (Fig. 4Aa-b), reflecting the high conformational flexibility of this region in the TYK2-KD simulated here (Fig. 2). Notably, the Com-A1 region is constrained by pKD in the autoinhibited form of JAKs, thus preventing relative movements of the N- and C-lobes necessary for kinase activity¹⁷, whereas it is released in mutation-activated JAK forms¹⁸. Taken together, these observations in conjunction with the important role of β_3 in ATP-binding (bears the KE-lysine, K930 and the C-spine residue, A928; Fig. S1), suggest an also important contribution of the dynamics of the β_2 - β_3 bulge in the dynamic assembly of the N-lobe β -structure upon ATP-binding, in TYK2/JAKs. Supporting this idea, *Tyk2* missense mutations that result in altering residues of dynamic nature in this region, have been also identified in cancer patients (e.g. G922D/S, T923I⁵²).

Com-A2 (in dark red) is TYK2/JAK-specific and comprises the β_4 - β_5 connecting loop, which is placed on top of the Com-A β -structure (Fig. 4Ab) and is also highly dynamic in the TYK2-KD simulated here (Fig. 2). Com-A2 is a small TYK2 community, which is however, allosterically coupled with Com-A during the analyzed time-frame of the simulation (Fig. 4Ac), implying an equally important role for the dynamics of the β_4 - β_5 connecting loop in correct ATP-positioning.

Com-N (in dark grey) is a TYK2/JAK-specific community and includes the N-terminus (aa: P889 to L897) of the TYK2-KD simulated here (Fig. 4Aa-b), which actually corresponds to the C-terminal end of the pKD-KD connecting region of JAKs¹⁷. Com-N is a solely MC-membered community (Fig. S7), which is however dynamically coupled with two other N-lobe communities, -A and -A1 (Fig. 4Ac), also implying a role of Com-N forming residues in the dynamic assembly of the N-lobe β -structure upon ATP-binding and its correct, catalytically competent positioning. Indeed, this TYK2 region is also a cancer hotspot⁵².

The Regulatory, Com-C

Com-C (in yellow) corresponds to the regulatory community, which is involved in the dynamic assembly of the R-spine upon AL phosphorylation and in the TYK2-KD simulated here, it comprises most of the regulatory C-helix (Fig. 4Aa; *MC*) including the R-spine residue, L951 and the glutamate residue, E947 of the KE salt-bridge. The SC-membered Com-C (Fig. 4A; *SC*), on the other hand, emerged as the dominant (Fig. S7) and most interconnecting SC-community in the simulated TYK2-KD (Figs. 4Ac; *SC* and S6; *bar at "C"*). Indeed, and in line with its role in the dynamic R-spine assembly, the TYK2 Com-C^{SC} integrates almost all the R-spine residues (yellow side-chains in Figs. 4B/ "*R spine*" and S8) and forms an extended positively-correlated side-chain communication network that dynamically connects α C with the assembled phosphorylated AL and its neighboring C-terminal and N-terminal halves of helices α E and α F (yellow in Fig. 4Aa-b; *SC*).

The highly correlated Com-C side-chain motions detected here, are reminiscent of synchronous side-chain movements observed in the inner hydrophobic core of PKA in response to ATP-binding and are proposed to serve dynamics-driven allosteric activation of EPKs². Indeed, additionally to R-spine forming residues, the side-chains of the gatekeeper, M978 (see also Com-A description above) and of other hydrophobic residues of the TYK2 inner hydrophobic core, predicted to serve dynamic bridging of the spines in activated EPKs, are part of the identified TYK2 Com-C^{SC} (yellow SC bars in Fig. 5B/ "*Bridging the Spines*"). Validating these findings, replacement, e.g. of I950 or of the gatekeeper, M978 by a valine (shorter) residue, results in uterine or colorectal cancer, respectively⁵². Based on this notion, additional Com-C^{SC} members with

long hydrophobic side-chains, emerged from this study as potential contributors to the dynamic assembly of the R-spine and to dynamically bridging the spines of the TYK2-KD during the activation process (Fig. 5B/”*Hydrophobic Support to the R-spine*”).

In total, Com-C integrates transiently key residues for intra-KD signal transduction upon ATP/Mg binding and AL phosphorylation and emerged as a key player in allosteric regulation of TYK2 activity.

C-lobe Communities:

Com-D (in dark green) corresponds to the catalytic community and its structurally-equivalent in PKA, for example, is involved in ATP-binding and proper positioning of the γ -phosphate and magnesium ions for phosphoryl-transfer to the substrate as well as in the assembly of the C-spine³². In the TYK2-KD simulated here, however, Com-D emerged as a solely MC-community (Fig. S7), most probably due to the absence of a protein-substrate and/or of a second magnesium ion in the simulation. Supporting a catalytic role, however, the TYK2 Com-D comprises the main-chains of most of the catalytic residues and their neighboring amino-acids involved in substrate recognition and the assembly of the catalytic spine, such as the D-helix, the catalytic loop (¹⁰²⁵AARNV¹⁰²⁹) and its neighboring C-terminal portion of the α F-helix (green in Fig. 4Aa-b; *MC*). In total, Com-D provides the backbone atoms of four out of six C-spine residues of the C-lobe (Fig. 4B/”*C-spine*”), and is allosterically coupled with both N-lobe and C-lobe communities with the strongest couplings (thick edges in Fig. 4Ac) being with Com-A and Com-FG (described below). In addition, when negatively correlated motions are taken into account, several Com-D side-chains are merged with Com-FG (magenta Fig. 5B; *SC*) and these observations are in agreement with literature data showing that the D-helix and its neighboring substrate-binding regions are more dynamic in Tyr-specific kinases.⁵⁷

Com-E (in blue) is the second largest community identified in the TYK2-KD simulated here (Fig. S7) and comprises most of the E-helix, the β 7- β 8 region that is directly linked to the catalytic- and MG-loops, and the α C- β 4 linker (Fig. 4Ab; *MC*). Indeed, the two latter regions have been reported to move as a rigid block serving as pivot point for C-helix relative movements in protein kinases⁵⁷. Indeed, Com-E provides the main-chains of several spine-bridging residues as well as the side-chains of almost all the C-lobe C-spine-forming residues (blue bars in Fig. 4B). It also comprises the side-chains of several residues that contribute to sensing ATP-binding as well as residues in the vicinity of the substrate-binding region, such as the D-helix and the C-terminal half of the α F-helix (blue in Fig. 5Aa; *SC*).

Moreover, Com-E appears to be the only TYK2 community that mediates allosteric coupling of most of the C-lobe communities with the ATP-binding (Com-A) and the R-spine (Com-C) communities, in both the MC- and SC-based analyses (see edges of E spheres, in Fig. 4Ac) and emerged as the third most bridging TYK2 SC-community (Fig. S6). These observations further support a role of the TYK2 Com-E as pivot community to serving allosteric communication during relative N-/C-lobe movements.

Combined these findings, indicate that Com-E represents the C-spine assembly stabilizing community. Based on this idea, several additional Com-E^{SC} hydrophobic residues, emerged from our analysis as potential contributors to the dynamic assembly of the C-spine in TYK2 (Fig. 4B/”*Hydrophobic Support to the C-spine*”).

Com-F (in brown) corresponds to the activation community in other EPKs and is the largest community of positively-correlated backbone atoms identified in the simulated TYK2-KD of this study (Fig. S7). Com-F^{MC} is built around the activation loop, including the β 6 β 9- and β 10 β 11-sheets (bearing Y1054p-Y1055p) that are hallmarks of assembled activation loops (extended, AL-OUT) in active PTKs⁴, strongly supporting its role as the TYK2 activation community (brown in Fig. 4Aa-b; *MC*). It also comprises the half of the α F-helix that provides a scaffold for the R-spine assembly (via D1083), as well as the C-terminal end of the α E-helix preceding β 6 (brown in Fig. 4Aa-b; *MC*). Interestingly, Com-F^{MC} is allosterically coupled with almost all the C-lobe communities as well as with the N-lobe Com-B and the regulatory Com-C, communities (Fig. 4Ac; *MC*). Indeed, it is noteworthy that the TYK2 Com-F^{MC} comprises the main-chains of two residues of the C-helix; namely, W944 and K945 (brown stripe in Fig. 5Aa; *MC*), reflecting a strong allosteric connection between AL and the C-helix in the TYK2-KD simulated here. Since interactions

between the C-helix and AL are characteristic of active kinases,⁵⁷ this observation infers an important role of these amino-acids in allosteric TYK2 activation.

The SC-membered Com-F, on the other hand, is the third largest SC-community (Fig. S7) and mostly a “bridging” community (Fig. S6; bar at “F”), reflecting the highly fluctuating phosphorylated AL in the TYK2-KD simulated here (see also Fig. 2). Interestingly, the TYK2 Com-F^{SC} comprises the side-chains of the entire FG-helix (Fig. 4Aa; *SC*) and therefore, it can also be regarded as the TYK2 Com-FG^{SC}. Notably, Com-F^{SC} is also strongly allosterically coupled with the R-spine (C^{SC}) and C-spine (E^{SC}) communities (Fig. 4Ac) and this observation combined with the Com-F^{MC} results, suggests a central role of α FG side-chains in dynamically bridging the spines and in dynamics-driven allosteric activation of TYK2.

Com-FG (in magenta) was named after the FG-helix as it includes the main-chain atoms of the entire α FG (Fig. 4Aa-b; *MC*) and is therefore, the major JAK-specific community. Interestingly, Com-FG extends to the ¹⁰⁶⁴PVFWY¹⁰⁶⁸ portion of the P+1 loop that determines Tyr-specificity for protein substrates and to its adjacent APE-helix (Fig. 4Aa; *MC*). Proline P1064 in particular, together with tryptophan W1067, is conserved in all PTKs and provides the interaction platform for the p-site tyrosine of protein substrates during catalysis, with also important role in auto-activation⁴.

Furthermore, Com-FG is also a solely main-chain community, which is allosterically coupled with the catalytic community, Com-D and the activation community, Com-F^{MC} with the strongest coupling being with Com-D (Fig. 4A.c; *MC*), supporting a central role of the FG-helix, in dynamically coupling substrate specificity with catalysis.

Taken together these findings, strongly suggest that the functional role of the TYK2 Com-FG, is related to determining TYK2/JAKs-specificity for protein substrates and that the α FG and APE helices hold a central role in this, on top of the ¹⁰⁶⁴PVFW¹⁰⁶⁷ region of the P+1 loop.

Com-G (in pink) comprises the backbone atoms of the G-helix and of its preceding loop that connects it with the FG-helix (Fig. 4Aa-b; *MC*). This region, contributes to substrate binding as well as to binding regulatory proteins in JAKs, e.g. the SOCs proteins that inhibit JAK/STAT signaling⁵⁶. Com-G is a dynamic, also solely MC-community, which is however allosterically coupled with two other protein-protein interaction communities; namely, Com-FG and Com-H (Fig. 4Ac; *MC*). These observations reflect a dynamic character of the G-helix and especially of the loop that connects it with α FG (see also Fig. 2) and further support the notion that higher flexibility of substrate-binding in PTKs⁵⁷.

The GHI community *Com-H^{PKA}*, which is proposed to serve as docking site for regulatory protein-protein interactions in EPKs³², is split in two discrete communities in the TYK2-KD simulated here; namely, Com-H and Com-I:

Com-H (in purple) is the largest community of the TYK2 GHI helical subdomain (Fig. S7) and encompasses the α H helix as well as its flanking linkers with the α G (α G- α H) and α I (α H- α I) helices (Fig. 4Aa-b; *MC*). Interestingly, this region has been suggested to act as scaffold for trans-activation of homo- or hetero-dimeric protein kinases in an asymmetric-dimer mode⁵⁹. Given the homo-/hetero-meric mode of action of JAKs, this concept points to the TYK2 Com-H as a potential dimerization and/or scaffolding community. Notably, however, the side-chains of the α H- α I linker split off from Com-H and are merged with the neighboring community Com-I, reflecting the dynamic character of this region in the simulated TYK2-KD (see also §3.2 and §3.3; and Figs. 3 and 4). Combined these results, imply that the dynamics of the α H- α I linker is most likely crucial for preventing unspecific scaffolding.

Com I (in dark blue) comprises the main-chains of almost the entire α I (Fig. 4Aa-b; *MC*) and is dynamically coupled with the C-lobe communities, -F, -E and -H (Fig. 4c; *MC*), implying involvement of the TYK2 α I-helix in allosteric TYK2 activation. In line with this hypothesis, in the autoinhibited form, the α I helix, along with its spatial neighboring helix, α E (see Fig. 4Ab; *MC/back*), is constrained by other TYK2 sub-domains (based on¹⁸). More importantly, the side-chains of the Com-I^{SC} residue members of the α H- α I connecting region, bridge it with Com-H (Fig. S6) and form a dynamic surface underneath the R-spine community,

Com-C^{SC} (dark blue and yellow-surfaces, respectively in Fig. 4Ab; *SC/back*). These results, combined with the observation that Com-I^{SC} serves allosteric coupling of Com-C and Com-H communities (Fig. 4Ac; *SC*), suggest a key role of the dynamics of the Com-I side-chains, and especially of the α H- α I linker, in allosteric regulation of TYK2 activity.

Com-P+1 (in black in Fig. 4b; *MC*) encompasses the portion of the P+1 loop (aa: R1058-S1063) that is involved in binding protein-substrate positions [?] p+1 (as deduced by the substrate-mimicking binding of SOCs⁵⁶). It is also a solely MC-community but allosterically coupled with Com-F^{MC} (Fig. 4A.c), suggesting an important role of the P+1 loop dynamics in coupling activation with substrate binding.

In total, the community analysis suggested that the TYK2-KD simulated here is in a dynamically committed state for substrate recognition/binding and catalysis; and pointed to several amino-acids as potential mediators of dynamic-driven allosteric communications within the TYK2-KD at this step of the catalytic cycle.

3.4.2. Identification of amino-acids serving dynamics-driven allosteric communications within the simulated TYK2-KD

To identify amino-acids that contribute the most to the TYK2-KD allosteric community network identified in this study (Fig. 4c), side-chain betweenness centrality values, as obtained by Bio3D (see *Methods*), were extracted and plotted along the amino-acid sequence. As shown in Fig. 5, several amino-acids across the TYK2-KD primary structure are characterized by higher centrality values of their side-chains, suggesting a greater role in mediating allosteric signal propagation within the catalytic domain and in dynamically coupling the spines as a result of ATP.1Mg²⁺-binding and AL phosphorylation. Indeed, among these amino-acids are C-spine (V911, A928, T1090) and R-spine (H1021, D1083) residues, conserved hydrophobic amino-acids dynamically bridging the spines in activated kinases (L976, M978) as well as catalytically important residues (N1028) and substrate binding amino-acids, including the PTK-characteristic¹⁰⁶⁴PVFW¹⁰⁶⁷ portion of the P+1 loop (Fig. 5). Notably, most of these residues (more than 40%) are associated with cancer-related TYK2 amino-acid substitutions, further supporting the functional relevance of our findings. Interestingly, among these are three α FG side-chains, including P1104 (Fig. 5), supporting a pivotal role of the FG-helix in dynamics-driven allosteric feed-back to the active site at this step of the catalytic cycle.

To elaborate on this in the light of the community map analysis results, amino-acids with higher side-chain centrality values ($Bc^{sc} > \text{average } Bc^{sc}=356$; labelled in Fig. 5) were categorized according to their corresponding communities and along with the predicted functional community roles (§3.4.1), are shown in Table I.

As shown in Table I and illustrated in Fig. 6, the higher centrality side-chains are organized into community groups with distinct identifiable functional contributions to the allosteric network of amino-acids and structurally, form interconnected dynamic motifs that wrap around the spines, thus contributing to their dynamic assembly and coupling in response to ATP.Mg²⁺-binding and AL phosphorylation (compare Figs. 6 and S8). We then sought to clarify the functional significance of these amino-acids in connection with sequence conservation and involvement in cancer (collectively shown in Table II).

Functional relevance of the dynamically coupled allosteric communication network of amino-acids identified in this study

Intra-KD allosteric signaling related to ATP-adenine positioning (Com-A^{SC})

A functional group of side-chains strongly contributing to the allosteric communication network of amino-acids identified in this study, is Com-A^{SC} (Table I/*row-A*; red in Fig. 6A); i.e. it is related to correct positioning of the ATP-adenine ring. Indeed, it provides side-chains of several conserved and well-characterized ATP-binding related amino-acids, as already mentioned, such as V911 and A928 (C-spine-forming), L976 (spines-bridging), E979 (ATP-adenine-binding) as well as conserved glycine residues, G904 and G909 of the G-loop (Table I and Fig. 6). As expected, and further validating the functional relevance of these results, this functional group is a cancer hotspot (Table II; “*ATP-adenine-positioning Block*”).

Notably, an additional, unique to TYK2 side-chain emerged from our analysis as potential contributor to mediating intra-KD allosteric communications upon ATP-binding; namely, S912, which is a glutamate in the other JAKs (Table II). This residue follows amino-acid V911 and projects upwards the N-lobe β -structure (Fig. 6B/*right*), indicating an important role in controlling its dynamic assembly and correct shaping of the ATP-binding pocket. Indeed, the observed fluctuations of the N-lobe β -structure would interfere with pKD in the autoinhibited full-length TYK2 (Fig. 6B). In addition, and interestingly enough, the N-lobe β -structure has been recently reported to be involved in transient interactions with pKD during transactivation of JAKs⁶⁰, further supporting its role in assisting dynamically coupled intra-KD allosteric communications during the activation process, predicted by our analysis. Further corroborating our findings, substitutions of S912 affecting the dynamics (S912I) or electrostatic interactions (S912R) in the region, along with a splice-site mutation of residue G970 of the β 4- β 5 loop (Com-A2; Fig. 6B/*right*), are found in several cancers (Table II).

Intra-KD allosteric signaling related to kinase activity regulation; dynamic assembly of the R-spine/activation (Com-C^{SC})

The predominant functional group of strongly signaling side-chains (bold in Table I; Fig. 6A), is related to the dynamic assembly of the R-spine (*C-row*; yellow-surfaces), with the sub-group connected to activation being the most populated (Table II; “*Activation block*”). Indeed, the latter sub-group includes the R-spine residue, D1083 and its adjacent hydrophobic residue, V1084 of α F as well as the hydrophobic amino-acids I1020, A1045 and V1048 (Fig. 6A) of the β 6 β 9-sheet, the assembly of which is a hallmark of activated (phosphorylated AL-OUT) kinases.

Notably, the strongly signaling Com-C^{SC} members also include the PTK-specific tryptophan residue, W1067 of the P+1 loop as well as a distal phenylalanine located in α I, F1162 (Fig. 6A), indicating a key role of these side-chains in allosteric signal transmission to the active site during TYK2 activation. Supporting this idea, a splice-site mutation of W1067 is a cancer hotspot (Table II). Given that, in addition to W1067, the phenylalanine at position 1162 is also conserved in PTKs (data not shown), these results imply that the dynamics-driven feed-back of these side-chains to the active site observed here, may also hold true for other JAKs/PTKs.

Interestingly, the C-spine residue T1090, also emerged as a significant contributor to allosteric signaling for the dynamic assembly of the R-spine (bold in Table I; Fig. 6A). This finding combined with the observation that T1090 shares the same functional group with the conserved Mg-binding asparagine, N1028 (Table II; “*Catalytic block*”), strongly suggests a pivotal role of T1090 in allosteric regulation of catalytic activity in response to ATP/Mg-binding. Corroborating this hypothesis, T1090 replacement by an isoleucine, which according to our analysis, is predicted to stabilize the C-spine (and/or the R-spine) and to result in an unregulated TYK2 catalytic activity, is found in glioblastoma (Table II). It is also noteworthy that T1090, is conserved only in TYK2 and its substrate, JAK1, whereas it preserves a more hydrophobic character in the other two JAK-members (Table II) and in PTKs (data not shown), implying that the regulatory role of T1090 inferred for TYK2, may also apply to JAK1 and is therefore, important to coordinating cytokine signaling.

The Com-C^{SC} functional group provides additional side-chains (yellow in Fig. 6B) that further contribute to the allosteric communication network. These include the hydrophobic side-chains of residues, M978 (gatekeeper; see also §3.4: Com-A and -C), W944 (α C), M1011 and A1016 (α E), L1024 (catalytic loop), L1044 (DFG+1), A1047 (β 9), C1072 (APE+1), and A1081 (α F) as well as several polar amino-acids, including the R-spine, HRD-histidine, H1021 (Table I/*C-row*).

Interestingly, many of the aforementioned amino-acids, especially those of the activation block, which is a cancer hotspot (Table II; “*Activation block*”), are also either TYK2-unique or shared between TYK2 and its counterparts in cytokine signaling, JAK1 and JAK2. For example, besides T1090, amino-acids A1047 (of the activation loop) and C1072 (of the APE-helix) are also shared only by TYK2 and JAK1 among the JAKs, whereas the α C residue, W944 is TYK2-unique (Table II). W944 in particular, as already mentioned,

is predicted to contribute to allosteric TYK2 activation by dynamically linking α C with the activation loop (see §3.4.1; *Com-F*) and indeed, it is also mutated in cancer (Table II). Another TYK2-unique amino-acid of the activation block is A1016, which is a serine/threonine in the other JAKs and interestingly, it is replaced by threonine or valine in several cancers (Table II), further supporting its predicted role in dynamics-driven allosteric TYK2 activation. Finally, residue A1081, which in the TYK2-KD simulated here, was found to lose contact with A1156 at the α H- α I loop (§3.3 and Fig. S4), also emerged as part of the activation block (Table II), implying an important role of this amino-acid in dynamics-driven allosteric TYK2 activation. Indeed, replacement of this side-chain by the longer hydrophobic side-chain of valine, as identified in colorectal cancer (Table II), is therefore expected to stabilize the network and to result in an unregulated activation. These results combined with the observation that A1081 is shared between TYK2 and its substrates JAK1 and JAK2 (Table II), support the notion that the dynamics of its α H- α I connecting region not only holds a crucial role in dynamics-driven allosteric activation of TYK2 but it most probably dictates recognition of its counterparts in cytokine signaling, implying a role in coordinating specific cytokine responses.

Taken together these findings, indicate a TYK2-unique amino-acid composition of intra-KD allosteric signaling related to its activation in response to ATP.1Mg²⁺-binding and/or AL-phosphorylation.

In total, the Com-C^{SC} correlated motions observed here, exemplify the contribution of AL phosphorylation to the dynamic assembly of both the R-spine and the catalytic site, and suggest that AL-phosphorylated TYK2 is dynamically committed to substrate recognition/binding and catalysis in response to ATP.1Mg²⁺-binding. This in turn suggests that AL phosphorylation and/or ATP.1Mg²⁺-binding may bypass cytokine-dependent hetero-dimerization and may account for its distinct activity.

Intra-KD allosteric signaling related to the dynamic assembly of the catalytic-spine/catalysis (Com-E^{SC})

The second most contributing signaling group of side-chains is that related to the dynamic assembly of the catalytic, C-spine (Table I/*E-row*; Fig. 6), with the most prominent contributions being provided by two functional amino-acid blocks; namely, the C-spine and the C-spine-stabilizing blocks (Table II).

The functional relevance of the C-spine block, in particular is corroborated by the observation that it is entirely a cancer hotspot with amino-acids A1025-A1026 of the catalytic loop, being the strongest contributors to allosteric intra-KD signaling (bold in Table I; Table II). This group is conserved in the JAKs, with the notable exception of amino-acid at position 996, which is a glycine uniquely in TYK2 (Table II). G996 is located at the C-terminal-end of the α D- α E connecting loop, which is shorter in TYK2 compared to the other JAKs (Fig. S5), implying an important role of the dynamics of this TYK2 region in catalytic activity. Moreover, 50% of cancer-related mutations associated to this group, result in longer hydrophobic side-chains (A1026V, L990M, T096M; Table II) that according to our analysis, are predicted to stabilize the C-spine assembly, further supporting the importance of this amino-acid block in dynamics-driven allosteric control of catalytic activity.

The C-spine-stabilizing block of dynamically coupled amino-acids (Table II; “*C-spine-stabilizing block*”), on the other hand, includes residues of the N-terminal half of the α E helix and of the β 7-strand preceding the Mg-binding loop that also map in the interface with other TYK2-domains in the autoinhibited form (blue surfaces in Fig. 6B/*right*). Among these amino-acids are, V1037 and other hydrophobic side-chains identified here as contributors to the dynamic assembly of the C-spine (Fig. 4B; “*Hydrophobic support to the C-spine*”) as well as the conserved K1038, which signals through its interaction with the ATP-adenine-bound, E979 (Fig. 6B/*right*). Corroborating the functional significance of this group, an isoleucine substitution for V1037, expected to stabilize the C-spine and to disrupt the dynamic bridging of the spines, leads to glioblastoma (Table II). Notably, several amino-acids of this block are also either TYK2-unique (L1036, A998, L983) or shared with JAK1 (A1004, Q999) (Table II). Interestingly, residue Q999 is an arginine in JAK3 and an arginine substitution for this TYK2 glutamine has been identified in colorectal cancer patients (Table II). Taken together, the results of this group point to a unique interface of the TYK2-KD with its other sub-domains and imply an important contribution of the dynamics of this region to allosteric signaling for substrate recognition and catalysis.

Intra-KD allosteric signaling related to substrate binding (Com-F^{SC});

Role of aFG in allosteric signal transmission; TYK2-specificity

As already mentioned, three side-chains of the TYK2 FG-helix; namely, P1104, P1105 and I1112, emerged from our analysis as strong mediators of allosteric signal propagation to the active site (bold in Table I, Fig. 6A).

As shown in Table I (*F-row*), these side-chains are part of the second most contributing functional group of dynamically signaling side-chains (brown in Fig. 6A), and of the TYK2-specificity-determining block, in particular (Com-FG^{MC}/F^{SC} cell in Tables I, II). The functional role of the latter derived from the observation that the conservation pattern of this amino-acid block is JAK-specific but in a TYK2-unique combination; namely, the strongly signaling triad of V1065, P1105 and I1112 side-chains of this group, is TYK2-unique (Table II; “*TYK2-specificity determining block*”). Corroborating these findings, side-chain substitutions affecting the specificity-signaling of P1105 and I1112, along with the P1104 to Ala change, are associated with cancer (Table II). Further supporting the role of this amino-acid group as specificity-determinant, P1104A was found to exhibit altered specificity in IFN-stimulated cells²⁷ and an Ala substitution for an isoleucine equivalent to I1112 (I1065^{JAK2}), has been reported to prevent (auto)phosphorylation of JAK2²⁴. Notably, residues A1069-P1070, are also part of the TYK2-specificity-determining block of side-chains identified in this study (Table II; “*TYK2-specificity determining block*”). This indicates an important role of the APE-motif in determining TYK2 specificity, and this may be facilitated by the disruption of the APE-R salt-bridge observed in this study (§3.2). Supporting this idea, a Gly substitution for A1069, expected to disrupt TYK2-specificity signaling, is associated with lung adenocarcinoma (Table II).

Furthermore, P1104 and its adjacent P1105, are part to a central hub (encircled in Fig. 6B) of the identified allosteric communication network of amino-acids that serves allosteric coupling of substrate-binding with the catalytic machinery. More specifically, and as shown in Fig. 7A, the P1104 side-chain ring packs against the W1067 indole-ring, which in turn is hydrogen-bonded to the conserved α F glutamate, E1093, whereas P1105 packs in a ring-stacking interaction with the F1066 phenol ring; and all these dynamically coupled interactions were preserved during the analyzed time-frame of the MD simulation, as indicated by narrow distributions and small values (< 5.5 Å) of the corresponding side-chain distances (Fig. S9). The side-chain of E1093 in turn, is anchored to the catalytic loop through a hydrogen-bond with the main-chain amide-group of A1026 (Fig. 7A) and this interaction also persisted (distance < 4 Å; Fig. S9). Interestingly, the PTK conserved proline residue, P1064 is transiently involved in Van-der-Waals interactions with the W1067 (as reflected by the two-peak distribution of their side-chain distance; Fig. S9) and this way, it also dynamically feeds-back to the catalytic machinery.

Taken together, these findings support a pivotal role of the FG-helix in determining specificity and in allosteric intra-KD signal propagation for coupling substrate specificity/binding with catalysis.

Τηρ γατεκεεπερ Μ978 ανδ β3-λψσινε, Κ930

Several additional long-range interactions, mediated by regulatory signaling side-chains identified here (yellow in Fig. 7A), facilitate dynamics-driven communication of substrate-binding and ATP-binding sites (e.g. at K930). Part of this network is the gatekeeper, M978 (Fig. 7A), as already mentioned above. This is better illustrated in Fig. 7B, depicting a conformational ensemble in the analyzed simulation time-period. This involves oscillation of M978 and of other side-chains of the inner hydrophobic core, including R-spine-forming residues and the β 3 ATP-binding lysine, K930 (of the KE salt-bridge; see also Fig. 1), between conformations pointing either towards the ATP molecule or the C-helix (Fig. 7B); actually the M978 and K930 side-chain movements are negatively correlated in the TYK2-KD simulated here. This exemplifies the role of methionine M978 as the “gatekeeper” and as a regulatory switch in TYK2, as described in §3.4.1 (see Com-A and Com-C description). These findings are in perfect agreement with the deleterious effect of both K930 and M978 substitutions linked to catalytically impaired TYK2 variants or to cancer-associated amino-acid substitutions (e.g. Lys930 to Arg²⁷ or Met978 to Phe⁶⁰). Moreover, the region corresponding to Com-N identified in this study (e.g. aa: P889-T890), assists to shielding the entrance to the hydrophobic

core by transiently capping the hydrophobic pocket near the ATP-binding site (Fig. 7A). Notably, and on top of the regulatory side-chains, M978 and Y962, amino-acids P889 and T890 are also conserved in JAKs (Fig. S5) but not in other PTKs (data not shown), suggesting that the dynamically coupled capping network identified here and involving residues, P889-T890-Y962-M978-K930, may be JAK-specific. Validating the functional significance of this finding, and in addition to M978 and Y962, P889 and T890 are also mutated in cancer⁵². Moreover, and further supporting these results, this capping would be served by pKD in its activating position, in a mutation-activated full-length protein (based on¹⁸).

Notably, the ensemble of dynamic conformations observed here, involves the side-chain of the adjacent to A1026 in the catalytic loop, arginine R1027, which also switches between conformations either transiently packed against W1067 (cation- π interaction) or at a hydrogen-bonding distance with the ATP γ -phosphate, while remaining hydrogen-bonded to the HRD-aspartate, D1023, in the simulated TYK2-KD (Figs. 7B and S7). This R1027-mediated interaction network helps neutralizing the buildup of negative charge on the γ -phosphate and D1023, thus most probably preventing ATP hydrolysis in the absence of substrate. Indeed, R1027 does not contribute to allosteric signaling to the active site in the TYK2-KD simulated here, as reflected by the zero Bc^{SC} value of its side-chain (Fig. 5). Supporting these findings, replacement of R1027 by a histidine, results in tumor-associated TYK2 activation⁶¹.

Combined these observations suggest that the dynamic conformational ensemble we observed in this study reflects the ensemble of states existing prior to phosphoryl-transfer and serving dynamics-driven regulation of TYK2 catalytic activity prior to substrate binding and securing dynamic allosteric coupling of ATP-binding and activation with substrate recognition/binding and catalysis.

4. Conclusions

In conclusion, this work highlights the importance of conformational dynamics of the TYK2-KD in allosteric regulation of TYK2 kinase activity and is in perfect agreement with the notion that ATP-binding and conformational flexibility of the catalytic domains of JAKs are critical for their (both cytokine-dependent or mutation-driven) activation.^{27,60}

In summary, this study by combining a long, microsecond-scale MD simulation with community network analysis, shed light on the dynamic profile of the TYK2-KD in response to ATP.1Mg²⁺-binding and AL phosphorylation (activation process) and identified a dynamically coupled allosteric communication network of amino-acids across the lobes, serving intra-KD signaling for dynamics-driven allosteric regulation of TYK2 activity, prior to substrate binding. Corroborating the functional significance of the identified allosteric network of side-chains, more than 40% of the key residues identified in this study, are associated with cancer-related *Tyk2* missense or splice-site mutations, as revealed by screening of related databases.

In particular, our community analysis supports a pivotal role of amino-acids P1104, P1105 and I1112 of the FG-helix in the communication network and strongly suggests that the JAK-specific helical insert is not simply involved in substrate binding but it is actively implicated in determining TYK2 specificity and in dynamics-driven allosteric feed-back to the active site for coupling substrate specificity/binding with catalysis, thus substantiating its reported role in JAK activation.²⁴

In addition, the community-network analysis of this study was able to provide prediction on the functional consequences of cancer-associated TYK2 amino-acid substitutions in line with the notion that consideration of protein structural dynamics and identification of mediators of allosteric signaling, in particular, is advantageous to the identification of disease-causing mutations.⁶²

We propose that the pre phosphoryl-transfer step of the catalytic cycle studied here, is crucial for TYK2 activation and that the conformational dynamics at this step, coordinated by α FG, is crucial to serving molecular recognition and this combined with the TYK2-unique amino-acid composition of several regions, including the β 3- α C and α H- α I connecting regions, may account for its distinct specificities reported in the literature (e.g. for TYK2-specific sites of STAT3^{9,15} and/or for non-canonical, cytokine-independent TYK2 functions⁶¹).

In total, this work adds to knowledge towards an in-depth understanding of TYK2 activation and specificity determinants and we believe that the dynamic profile provided here, combined with the TYK2-unique amino-acid composition of the allosteric communication network identified in this study, may be valuable towards a structure-based design of allosteric TYK2-specific inhibitors.

References

1. Huse M, Kuriyan J. The conformational plasticity of protein kinases. *Cell*. 2002 May 3;109(3):275-82. doi: 10.1016/s0092-8674(02)00741-9.
2. Kim J, Ahuja LG, Chao FA, Xia Y, McClendon CL, Kornev AP, Taylor SS, Veglia G. A dynamic hydrophobic core orchestrates allostery in protein kinases. *Sci Adv*. 2017 Apr 7;3(4):e1600663. doi: 10.1126/sciadv.1600663.
3. Kannan NK, Neuwald AF. Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? *J Mol Biol*. 2005 Sep 2;351(5):956-72. doi: 10.1016/j.jmb.2005.06.057. PMID: 16051269.
4. al-Obeidi FA, Wu JJ, Lam KS. Protein tyrosine kinases: structure, substrate specificity, and drug discovery. *Biopolymers*. 1998;47(3):197-223. doi: 10.1002/(SICI)1097-0282(1998)47:3<197::AID-BIP2>3.0.CO;2-H.
5. Pellegrini S, Dusanter-Fourt I. The structure, regulation and function of the Janus kinases (JAKs) and the signal transducers and activators of transcription (STATs). *Eur J Biochem*. 1997 Sep 15;248(3):615-33. doi: 10.1111/j.1432-1033.1997.00615.x.
6. Leonard WJ, O'Shea JJ. Jaks and STATs: biological implications. *Annu Rev Immunol*. 1998;16:293-322. doi: 10.1146/annurev.immunol.16.1.293.
7. Kyogoku C, Tsuchiya N. A compass that points to lupus: genetic studies on type I interferon pathway. *Genes Immun*. 2007 Sep;8(6):445-55. doi: 10.1038/sj.gene.6364409.
8. Ghoreschi K, Laurence A, O'Shea JJ (2009). Janus kinases in immune cell signaling. *Immunol Rev* 2009; 228(1):273-287. doi: 10.1111/j.1600-065X.2008.00754.x.
9. Buchert M, Burns CJ, Ernst M (2015) Targeting JAK kinase in solid tumors: emerging opportunities and challenges. *Oncogene* 2015, 1-13. doi: 10.1038/onc.2015.150
10. Leitner NR, Witalisz-Siepracka A, Strobl B, Müller M. Tyrosine kinase 2 - Surveillant of tumours and bona fide oncogene. *Cytokine*. 2017 Jan;89:209-218. doi: 10.1016/j.cyto.2015.10.015.
11. Wöss K, Simonović N, Strobl B, Macho-Maschler S, Müller M. TYK2: An Upstream Kinase of STATs in Cancer. *Cancers (Basel)*. 2019 Nov 5;11(11):1728. doi: 10.3390/cancers11111728.
12. Karjalainen A, Shoebridge S, Krunić M, Simonović N, Tebb G, Macho-Maschler S, Strobl B, Müller M. TYK2 in Tumor Immunosurveillance. *Cancers (Basel)*. 2020 Jan 8;12(1):150. doi: 10.3390/cancers12010150.
13. Borcherdig DC, He K, Amin NV, Hirbe AC. TYK2 in Cancer Metastases: Genomic and Proteomic Discovery. *Cancers (Basel)*. 2021 Aug 19;13(16):4171. doi: 10.3390/cancers13164171.
14. Wan J, Fu AK, Ip FC, Ng HK, Hugon J, Page G, Wang JH, Lai KO, Wu Z, Ip NY. Tyk2/STAT3 signaling mediates beta-amyloid-induced neuronal cell death: implications in Alzheimer's disease. *J Neurosci*. 2010 May 19;30(20):6873-81. doi: 10.1523/JNEUROSCI.0519-10.2010.
15. Mori R, Wauman J, Icardi L, Van der Heyden J, De Cauwer L, Peelman F, De Bosscher K, Tavernier J. TYK2-induced phosphorylation of Y640 suppresses STAT3 transcriptional activity. *Sci Rep*. 2017 Nov 21;7(1):15919. doi: 10.1038/s41598-017-15912-6.
16. Wallweber HJ, Tam C, Franke Y, Starovasnik MA, Lupardus PJ. Structural basis of recognition of interferon- α receptor by tyrosine kinase 2. *Nat Struct Mol Biol*. 2014 May;21(5):443-8. doi: 10.1038/nsmb.2807. Epub 2014 Apr 6. PMID: 24704786; PMCID: PMC4161281
17. Lupardus PJ, Ultsch M, Wallweber H, Bir Kohli P, Johnson AR, Eigenbrot C. Structure of the pseudokinase-kinase domains from protein kinase TYK2 reveals a mechanism for Janus kinase (JAK) autoinhibition. *Proc Natl Acad Sci U S A*. 2014 Jun 3;111(22):8025-30. doi: 10.1073/pnas.1401180111.
18. Glassman CR, Tsutsumi N, Saxton RA, Lupardus PJ, Jude KM, Garcia KC. Structure of a Janus kinase cytokine receptor complex reveals the basis for dimeric activation. *Science*. 2022 Apr 8;376(6589):163-

169. doi: 10.1126/science.abn8933.
19. Tsui V, Gibbons P, Ultsch M, Mortara K, Chang C, Blair W, Pulk R, Stanley M, Starovasnik M, Williams D, Lamers M, Leonard P, Magnuson S, Liang J, Eigenbrot C. A new regulatory switch in a JAK protein kinase. *Proteins*. 2011 Feb;79(2):393-401. doi: 10.1002/prot.22889.
 20. Gauzzi MC, Velazquez L, McKendry R, Mogensen KE, Fellous M, Pellegrini S. Interferon-alpha-dependent activation of Tyk2 requires phosphorylation of positive regulatory tyrosines by another kinase. *J Biol Chem*. 1996 Aug 23;271(34):20494-500. doi: 10.1074/jbc.271.34.20494.
 21. Yang J, Wu J, Steichen JM, Kornev AP, Deal MS, Li S, Sankaran B, Woods VL Jr, Taylor SS. A conserved Glu-Arg salt bridge connects coevolved motifs that define the eukaryotic protein kinase fold. *J Mol Biol*. 2012 Jan 27;415(4):666-79. doi: 10.1016/j.jmb.2011.11.035.
 22. Kornev AP, Taylor SS, Ten Eyck LF. A helix scaffold for the assembly of active protein kinases. *Proc Natl Acad Sci U S A*. 2008 Sep 23;105(38):14377-82. doi: 10.1073/pnas.0807988105.
 23. Kornev AP, Taylor SS. Dynamics-Driven Allostery in Protein Kinases. *Trends Biochem Sci*. 2015 Nov;40(11):628-647. doi: 10.1016/j.tibs.2015.09.002.
 24. Haan C, Kroy DC, Wüller S, Sommer U, Nöcker T, Rolvering C, Behrmann I, Heinrich PC, Haan S. An unusual insertion in Jak2 is crucial for kinase activity and differentially affects cytokine responses. *J Immunol*. 2009 Mar 1;182(5):2969-77. doi: 10.4049/jimmunol.0800572.
 25. Kaminker JS, Zhang Y, Waugh A, Haverty PM, Peters B, Sebisanoovic D, Stinson J, Forrest WF, Bazan JF, Seshagiri S, Zhang Z. Distinguishing cancer-associated missense mutations from common polymorphisms. *Cancer Res*. 2007 Jan 15;67(2):465-73. doi: 10.1158/0008-5472.CAN-06-1736.
 26. Tomasson MH, Xiang Z, Walgren R, Zhao Y, Kasai Y, Miner T, Ries RE, Lubman O, Fremont DH, McLellan MD, Payton JE, Westervelt P, DiPersio JF, Link DC, Walter MJ, Graubert TA, Watson M, Baty J, Heath S, Shannon WD, Nagarajan R, Bloomfield CD, Mardis ER, Wilson RK, Ley TJ. Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with de novo acute myeloid leukemia. *Blood*. 2008 May 1;111(9):4797-808. doi: 10.1182/blood-2007-09-113027.
 27. Li Z, Gakovic M, Ragimbeau J, Eloranta ML, Rönnblom L, Michel F, Pellegrini S. Two rare disease-associated Tyk2 variants are catalytically impaired but signaling competent. *J Immunol*. 2013 Mar 1;190(5):2335-44. doi: 10.4049/jimmunol.1203118.
 28. Couturier N, Bucciarelli F, Nurtdinov RN, Debouverie M, Lebrun-Frenay C, Defer G, Moreau T, Confavreux C, Vukusic S, Cournu-Rebeix I, Goertsches RH, Zettl UK, Comabella M, Montalban X, Rieckmann P, Weber F, Müller-Myhsok B, Edan G, Fontaine B, Mars LT, Saoudi A, Oksenberg JR, Clanet M, Liblau RS, Brassat D. Tyrosine kinase 2 variant influences T lymphocyte polarization and multiple sclerosis susceptibility. *Brain*. 2011 Mar;134(Pt 3):693-703. doi: 10.1093/brain/awr010.
 29. Qin W, Godec A, Zhang X, Zhu C, Shao J, Tao Y, Bu X, Hirbe AC. TYK2 promotes malignant peripheral nerve sheath tumor progression through inhibition of cell death. *Cancer Med*. 2019 Sep;8(11):5232-5241. doi: 10.1002/cam4.2386.
 30. Contreras-Cubas, C., García-Ortiz, H., Velázquez-Cruz, R. et al. Catalytically Impaired TYK2 Variants are Protective Against Childhood- and Adult-Onset Systemic Lupus Erythematosus in Mexicans. *Sci Rep* 9, 12165 (2019). doi:10.1038/s41598-019-48451-3
 31. Lesgidou N, Eliopoulos E, Goulielmos GN, Vlassi M. Insights on the alteration of functionality of a tyrosine kinase 2 variant: a molecular dynamics study. *Bioinformatics*. 2018 Sep 1;34(17):i781-i786. doi: 10.1093/bioinformatics/bty556.
 32. McClendon CL, Kornev AP, Gilson MK, Taylor SS. Dynamic architecture of a protein kinase. *Proc Natl Acad Sci U S A*. 2014; 111(43):E4623-31. doi: 10.1073/pnas.1418402111
 33. Foda ZH, Shan Y, Kim ET, Shaw DE, Seeliger MA. A dynamically coupled allosteric network underlies binding cooperativity in Src kinase. *Nat Commun*. 2015 Jan 20;6:5939. doi: 10.1038/ncomms6939. PMID: 25600932
 34. Chopra N, Wales TE, Joseph RE, Boyken SE, Engen JR, Jernigan RL, Andreotti AH. Dynamic Allostery Mediated by a Conserved Tryptophan in the Tec Family Kinases. *PLoS Comput Biol*. 2016 Mar 24;12(3):e1004826. doi: 10.1371/journal.pcbi.1004826.
 35. Ahuja LG, Kornev AP, McClendon CL, Veglia G, Taylor SS. Mutation of a kinase allosteric node

uncouples dynamics linked to phosphotransfer. *Proc Natl Acad Sci U S A.* 2017 Feb 7;114(6):E931-E940. doi: 10.1073/pnas.1620667114.

36. Bao ZQ, Jacobsen DM, Young MA. Briefly bound to activate: How the transient binding of a second catalytic magnesium alters the structure and dynamics of the CDK2 protein kinase active site during catalysis. *Structure.* 2011 May 11; 19(5): 675–690. doi:10.1016/j.str.2011.02.016
37. Abraham MJ, Murtola T, Schulz R, Páll S, Smith J. C., Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers *SoftwareX* 1 (2015) pp. 19-25
38. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput.* 2008 Mar;4(3):435-47. doi: 10.1021/ct700301q. PMID: 26620784.
39. Liang J, Tsui V, Van Abbema A, Bao L, Barrett K, Beresini M, Berezhkovskiy L, Blair WS, Chang C, Driscoll J, Eigenbrot C, Ghilardi N, Gibbons P, Halladay J, Johnson A, Kohli PB, Lai Y, Liimatta M, Mantik P, Menghrajani K, Murray J, Sambrone A, Xiao Y, Shia S, Shin Y, Smith J, Sohn S, Stanley M, Ultsch M, Zhang B, Wu LC, Magnuson S. Lead identification of novel and selective TYK2 inhibitors. *Eur J Med Chem.* 2013 Sep;67:175-87. doi: 10.1016/j.ejmech.2013.03.070.
40. Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 18, 2714-2723.
41. Fensome A, Ambler CM, Arnold E, Banker ME, Brown MF, Chrencik J, Clark JD, Dowty ME, Efremov IV, Flick A, Gerstenberger BS, Gopalsamy A, Hayward MM, Hegen M, Hollingshead BD, Jussif J, Knafels JD, Limburg DC, Lin D, Lin TH, Pierce BS, Saiah E, Sharma R, Symanowicz PT, Telliez JB, Trujillo JI, Vajdos FF, Vincent F, Wan ZK, Xing L, Yang X, Yang X, Zhang L. Dual Inhibition of TYK2 and JAK1 for the Treatment of Autoimmune Diseases: Discovery of ((S)-2,2-Difluorocyclopropyl)((1R,5S)-3-(2-((1-methyl-1H-pyrazol-4-yl)amino)pyrimidin-4-yl)-3,8-diazabicyclo[3.2.1]octan-8-yl)methanone (PF-06700841). *J Med Chem.* 2018 Oct 11;61(19):8597-8612. doi: 10.1021/acs.jmedchem.8b00917.
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983; 79: 926–935. doi: 10.1063/1.445869
43. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* 2010 Jun;78(8):1950-8. doi: 10.1002/prot.22711.
44. Sellis D, Drosou V, Vlachakis D, Voukkalis N, Giannakouros T, Vlassi M. Phosphorylation of the arginine/serine repeats of lamin B receptor by SRPK1—insights from molecular dynamics simulations. *Biochim Biophys Acta.* 2012; 1820: 44–55. doi: 10.1016/j.bbagen.2011.10.010
45. Darden T, York D, Pedersen L. An N log(N) method for Ewald sums in large systems. *J Chem Phys.* 1993; 98: 10089–10092. doi: 10.1063/1.464397
46. Grant, BJ, Skjaerven, L, Yao, X-Q. The Bio3D packages for structural bioinformatics. *Protein Science.* 2021; 30: 20– 30. <https://doi.org/10.1002/pro.3923>
47. Grant BJ, Rodrigues AP, ElSawy KM, McCammon JA, Caves LS. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics.* 2006 Nov 1;22(21):2695-6. doi: 10.1093/bioinformatics/btl461. Epub 2006 Aug 29. PMID: 16940322
48. Girvan M and Newman M. E. J. Community structure in social and biological networks, *Proceedings of the National Academy of Sciences* Jun 2002, 99 (12) 7821-7826; DOI: 10.1073/pnas.122653799
49. Schrodinger, L. & DeLano, W., 2020. PyMOL The Molecular Graphics System v 2.1.0 Schrodinger, LLC;). Available at: <http://www.pymol.org/pymol>.
50. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7:539 doi:10.1038/msb.2011.75
51. Wang Y, Wang Q, Huang H, Huang W, Chen Y, McGarvey PB, Wu CH, Arighi CN, UniProt Consortium. A crowdsourcing open platform for literature curation in UniProt *Plos Biology.* 19(12):e3001464 (2021)

52. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012 May;2(5):401-4. doi: 10.1158/2159-8290.CD-12-0095
53. Shan Y, Eastwood MP, Zhang X, Kim ET, Arkhipov A, Dror RO, Jumper J, Kuriyan J, Shaw DE. Oncogenic Mutations Counteract Intrinsic Disorder in the EGFR Kinase and Promote Receptor Dimerization. *Cell* 149, 860–870, May 11, 2012.
54. Kornev AP, Haste NM, Taylor SS, Eyck LF. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A.* 2006 Nov 21;103(47):17783-8. doi: 10.1073/pnas.0607656103.
55. Deminoff SJ, Ramachandran V, Herman PK. Distal recognition sites in substrates are required for efficient phosphorylation by the cAMP-dependent protein kinase. *Genetics.* 2009 Jun;182(2):529-39. doi: 10.1534/genetics.109.102178. Epub 2009 Apr 13. PMID: 19364808; PMCID: PMC2691761.
56. Liao NPD, Laktyushin A, Lucet IS, Murphy JM, Yao S, Whitlock E, Callaghan K, Nicola NA, Kershaw NJ, Babon JJ. The molecular basis of JAK/STAT inhibition by SOCS1. *Nat Commun.* 2018 Apr 19;9(1):1558. doi: 10.1038/s41467-018-04013-1
57. Shudler M, Niv MY. BlockMaster: partitioning protein kinase structures using normal-mode analysis. *J Phys Chem A.* 2009 Jul 2;113(26):7528-34. doi: 10.1021/jp900885w.
58. Azam M, Seeliger MA, Gray NS, Kuriyan J, Daley GQ. Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat Struct Mol Biol.* 2008 Oct;15(10):1109-18. doi: 10.1038/nsmb.1486. Epub 2008 Sep 14. PMID: 18794843; PMCID: PMC2575426.
59. Jura N, Zhang X, Endres NF, Seeliger MA, Schindler T, Kuriyan J. Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Mol Cell.* 2011 Apr 8;42(1):9-22. doi: 10.1016/j.molcel.2011.03.004.
60. Caveney NA, Saxton RA, Waghay D, Glassman CR, Tsutsumi N, Hubbard SR, Garcia KC. Structural basis of Janus kinase trans-activation, *Cell Reports*, Volume 42, Issue 3,2023, 112201. doi:10.1016/j.celrep.2023.112201.
61. Sanda T, Tyner JW, Gutierrez A, Ngo VN, Glover J, Chang BH, Yost A, Ma W, Fleischman AG, Zhou W, Yang Y, Kleppe M, Ahn Y, Tatarek J, Kelliher MA, Neuberg DS, Levine RL, Moriggl R, Muller M, Gray NS, Jamieson CH, Weng AP, Staudt LM, Druker BJ, Look AT. TYK2-STAT1-BCL2 pathway dependence in T-cell acute lymphoblastic leukemia. *Cancer Discov.* 2013 May;3(5):564-77. doi: 10.1158/2159-8290.CD-12-0504. Epub 2013 Mar 7. PMID: 23471820; PMCID: PMC3651770.
62. Ponzoni L, Bahar I. Structural dynamics is a determinant of the functional significance of missense variants. *Proc Natl Acad Sci U S A.* 2018 Apr 17;115(16):4164-4169. doi: 10.1073/pnas.1715896115. Epub 2018 Apr 2. PMID: 29610305; PMCID: PMC5910821.

Table I. TYK2 amino-acids mediating dynamically coupled allosteric communications within the simulated TYK2-KD($Bc^{sc} > \text{average } Bc^{sc}=356$; labeled in Fig 5), categorized according to their corresponding communities identified in this study (as in Fig.4A). In bold: residues with statistically significant higher Bc^{sc} values ($Bc^{sc} > \text{average } Bc^{sc} + 1\sigma = 1000$), suggesting a pivotal role of their side chains in allosteric signal propagation within the TYK2-KD during the activation process. Potential functional roles predicted by the community analysis (details in § 3.4.1), are also included.

MC Com SC Com	Com-A	Com-A2	Com-B	Com-D	Com-E	Com-F	Com-FG	Com-G	Com-H
<i>Potential Functional Role</i>	<i>ATP-binding/positioning</i>	<i>Senses ATP-binding</i>	<i>Ασπιστς ποσιτιονινγ οφ τηε α^η ηελιξ</i>	<i>Catalytic; ATP/Mg-binding</i>	<i>C-spine Pivot community for αC helix and N-/C-lobe movements</i>	<i>Activation: (AL phosphorylation)</i>	<i>JAKs/TYK2 specificity determining; Substrate binding</i>	<i>Assists binding of substrate proteins; binding of regulatory proteins, e.g. SOCs</i>	<i>Protein-protein interactions; substrate docking/dimerization/scaffolding (putative)</i>
-A	G904* G909* V911^{3,5} S912* L913 A928 ^{3*} L976 ¹ E979 Y980	G970	G937* Q939*	N1028 T1090 3,5*	G1010^x M1011	W944* H1015 A1016* I1020 H1021 ^{2,6*} L1024 ⁴ G1043 L1044 A1045 A1047 V1048⁵ Y1080* A1081* D1083^{2*} V1084 S1086*	W1067⁶ C1072		
-A2									
-B									
-C (Regulatory; R-spine assembly)	M978 ^{1*}								

MC	Com-SC	Com-A	Com-A2	Com-B	Com-D	Com-E	Com-F	Com-FG	Com-G	Com-H
-E (<i>C-spine assembly stabilization</i>)					L990* G996* A1025* A1026* E1093 ^{6*} T1096*	L983 A998 Q999* L1001 L1002 A1004 Q1006 L1036 V1037^{4*} K1038	F1087 G1088*			C1151
-F (or FG) (<i>substrate binding</i>)							W1085 V1089	V1065 F1066⁶ A1069* P1070 P1104^{6*} P1105* I1112*	G1117 V1121 L1124 L1128	R1133
-H					Y1092					P1137 C1140 V1144 L1147* N1150 P1160
-I										
N/A	P982						K1077*	P1064*		

(¹) *R-spine shell*, (²) *R-spine* and (³) *C-spine* corresponding amino-acids (Kornev et al, 2008)²²;

(⁴), (⁵) *Dynamic residues found to move synchronously in the inner hydrophobic core of active EPKs in response to ATP or ATP/substrate-peptide binding, respectively* (Kim et al 2017)²;

(⁶) *part of a coupled allosteric network identified in another non-receptor tyrosine kinase (c-Src; Foda et al 2015)³³*;

(*) (*x*): *TYK2 amino-acids associated with cancer-related missense/splice-site mutations (retrieved from the cBioPortal for Cancer Genomics; Cerami et al 2012)⁵²*;

<https://www.cbioportal.org>)

N/A denotes uncorrelated side-chain movements

Table II. Collective listing of the allosteric communication network of TYK2 amino-acids identified in this study (as in Table I): Predicted functional roles; Amino-acid sequence conservation in the JAK family (see also Fig.S5); Involvement in cancer (*).

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
-A: ATP-binding; ασσεμβλψ οφ τηε N-λοβε β-στρυτυρε· correct positioning of the ATP-adenine ring	-A: ATP-adenine-positioning Block	G904		“	“	“	Protein Change G904D	Cancer Type Melano
		G909		“	“	“	G909D; G909S;	Renal Clear Cell Carcinoma; Uterine Endometrial
		V911 ^{3,5}		“	“	“		
		S912	S	E	E	E	S912I; S912R	Intrahepatic cholangiocarcinoma ; Lung Squamous Cell Carcinoma
		L913	L	L	M	L		
		A928 ³	A	A	A	A	A928V;A928T	Colorectal Glioblastoma
L976 ¹			“	“	“			
E979			“	“	“			
Y980	Y	Y	Y	Y	Y			

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
	-C:R-spine assembly	M978 ¹	M	M	M	M	M978V	Colorec
	<i>N/A: senses ATP binding (hinge region)</i>	P982	P	P	P	P		
-B: <i>Dynamic α-ηελίξ ποστιονινγ (β3-α^α)</i>	-B: <i>δντριβυτες το α^α, ΑΑ-ΟΥΤ ιντεραστιονς</i>	G937	G	G	T	G	G937A	Unknow primary
-A2: <i>σεινες ΑΠΙ βινδινγ (β4-β5)</i>		Q939 G970	Q G	N G	E G	D G	Q939S (X970-splice)	Colorec Extrask Myxoid Chon- drosarc Small Bowel I roendoc Extrahe atic Cholang carcino Glioblas
-D: <i>Catalytic; C-spine contribution; ΑΠΙγ/Μγ-βινδινγ συβστρατε ρεσογνι-τιον/βινδινγ</i>	-C: <i>R-spine assembly Catalytic Block</i>	T1090^{3,5}	T	T	V	V	T1090I	
	-E: <i>C-spine assembly C-spine Block</i>	N1028 A1025	A	A	A	A	A1025T	Extrahe Cholang carcino

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
		A1026	A	A	T	A	A1026V	Lung Squamous Cell Carcinoma
		L990		“	“	“	L990M	Melanoma
		G996	G	N	D	D	G996R	Colorectal
		E1093		“	“	“	E1093K	Cutaneous Melanoma
		T1096		“	“	“	T1096M	Mucinous Adenocarcinoma/Myxoid Fibrosarcoma
	<i>-H: PP interactions /Scaffolding (positively correlated with aH)</i>	Y1092	Y	H	Y	Y		
<i>-E: C-spine stabilization</i>	<i>-C: R-spine assembly</i>	G1010		“	“	“	(X1010-splice)	Lung Squamous Cell, Invasive Breast, Endometrioid Carcinomas
<i>πιοτ σομμυνηψ φορ α~ηελιξ & N/~-λοβε μοεμεντς senses auto-inhibitory domains</i>		M1011		“	“	“		

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)	
-E: <i>C-spine assembly/stabilization</i> C-spine-stabilizing Block -F: <i>Activation community (ασσεμβλψ οφ α^α ωιτη τηε αστιατιον σεντερ: β6β9- ε β10β11- σηεετς βψ Τψρ- πηοσπηορψλατεδ ΑΛ)</i>	-C: <i>R-spine assembly</i> Activation Block	V1037⁴	V	V	V	V	V1037I	<i>Glioblastoma</i>	
		Q1006		“	“	“			
		L1036	L	Q	R	H			
		L1001		“	“	“			
		L1002	L	K	Q	L			
		A1004	A	A	T	S			
		A998	A	K	I	S			
		Q999	Q	Q	K	R		Q999R	<i>Colorectal</i>
		K1038		“	“	“			
		L983	L	S	Y	S			
		V1048⁵	V	I	L	L			
		I1020	I	V	I	V			
		A1045	A	T	T	A			
		V1084		“	“	“			
		D1083²		“	“	“		D1083G/H/Y	<i>Colorectal</i>
G1043		“	“	“					
H1021 ²		“	“	“		H1021R	<i>Prostate</i>		
S1086		“	“	“		S1086F	<i>Melanoma</i>		
L1024 ⁴		“	“	“					
L1044		“	“	“					
A1081	A	A	A	A	Q	A1081V	<i>Colorectal</i>		

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
		Y1080	Y	I	V	R	Y1080C	<i>Esophagus; squamous</i>
		H1015 A1016	H A	G S	G T	G S	A1016T/V	<i>Esophagus; Bladder; Endometrial; Gastric Adenocarcinoma</i>
		W944	W	L	F	F	W944R	<i>Adenocarcinoma</i>
	-E: C-spine assembly	A1047 F1087	A	A “	V “	L “		
	- F(FG): substrate binding	G1088 V1089		“	“	“	G1088E	<i>Prostate</i>
	N/A=senses β10 (Υ1054) phosphorylation	W1085 K1077	K	“ K	“ K	“ I	K1077T	<i>Intrahepatic</i>
-FG: TYK2/JAK1 specificity determining community	- C: R-spine assembly	W1067		“	“	“	(X1067-splice)	<i>Cervical squamous; Pancreatic Adenocarcinoma</i>
		C1072	C	C	S	S		

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
	-F(FG): substrate binding TYK2-specificity Determining Block	V1065	V	V	I	I		
		F1066 A1069	F	F “	F “	F “	A1069G	<i>Lung adenocarcinoma</i>
		P1070 P1104		“ “	“ “	“ “	P1104A	<i>Melanoma Diffuse Large B-Cell Lymphoid Neoplasia</i>
		P1105	P	M	P	S	P1105L	<i>Lung Squamous Cell Carcinoma</i>
		I1112	I	I	I	M	I1112V	<i>Uterine Endometrioid Carcinoma</i>

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
	<i>N/A=senses Tyr-phosphosite of protein substrates</i>	P1064		“	“	“	P1064I	<i>Melanoma Unknown primary</i>
-G: Assists substrate binding; Binding of regulatory proteins (e.g. SOCs)	-F(FG): substrate binding (substrate positions $p > +1$)	G1117	G	G	G	D		
		V1121	V	V	V	L		
		L1124	L	L	L	L		
		L1128		“	“	“		
-H; PP interactions; substrate docking/dimerization/scaffolding (putative)	-E: C-spine assembly	C1151		“	“	“		
	-	R1133		“	“	“		
	F(FG): substrate binding							
	-H:	P1137		“	“	“		
		C1140		“	“	“		
		V1144	V	V	I	V		
		L1147	L	L	I	L	L1147M	<i>Melanoma</i>

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
<i>-I; PP regulatory interactions, including binding of other sub-domains in the auto-inhibited form (e.g. FERM)</i>	<i>-I: PP regulatory interactions</i>	N1150	N	K	E	L		
	<i>-C: R-spine assembly</i>	P1160 F1162	P	T “	P “	P “		
	<i>-E: C-spine assembly</i>	H1173	H	L	R	W		
	<i>-I: PP regulatory interactions; Substrate recognition (putative)</i>	L1169 P1167	L P	F E	V L	L P		

Predicted Functional Role (MC-Com in Table I)	Predicted Functional Role of Bridging Community (SC-Com in Table I)	TYK2 aa	TYK2	JAK1	JAK2	JAK3	Mutations identified in cancer patients (*)	Mutations identified in cancer patients (*)
		T1161	T	S	S	S	T1161A	<i>Lung Squamous Cell Carcinoma</i>

(*) retrieved from the *cBioPortal for Cancer Genomics* (Cerami et al 2012)⁵²; <https://www.cbioportal.org>

(“) indicates amino-acid conservation throughout the PTK family

(\$\$) For involvement of P1104A in additional types of cancer see e.g. reviews in (Leitner et al 2017¹⁰; Borcherding et al 2021¹³)

(¹) R-spine shell, (²) R-spine and (³) C-spine corresponding amino-acids (Kornev et al, 2008)²²;

(4), (5) Dynamic residues found to move synchronously in the inner hydrophobic core of active EPKs in response to ATP or ATP/substrate-peptide binding (Kim et al 2017)²;

In red: TYK2-unique amino-acids.

Figure Legends

Figure 1 . Monitoring of distances, hallmarks of active EPKs . Monitoring of conserved salt-bridges (KE and APE-R) and of distances of the KE-lysine, K930 with the ATP α - (left) and β -phosphates (right), along the MD trajectory.

Figure 2 . Backbone root-mean-square-fluctuation (RMSF) analysis of the MD trajectory in the time-frame of 2-3 μ s. Important secondary structure elements and motifs discussed in the text, are boxed and labelled.

Figure 3 : Cross-Correlation analysis of the MD trajectory (2-3 μ s) focused on highly fluctuating regions of the C-lobe (see Fig. 2). Residues with highly positively correlated (correlation coefficient [?] 0.8) motions of C α (denoted MC) and side-chain (denoted SC) atoms are interconnected with red lines on the corresponding representative snapshot of the MD simulation. Amino-acids discussed in the text are labelled.

Ect Figure 4: Community network analysis of the MD trajectory unravels the dynamic profile of the TYK2-KD in response to ATP.1Mg²⁺-binding and AL-phosphorylation. **(A)**Communities obtained by positive correlations-based community map analyses of the MD trajectory (2-3 μ s), are mapped on the representative MD snapshot and depicted as: **(a)** cylinder-cartoon and **(b)** surface representations. **(c)** 2D-graph illustration of the allosteric community network; The communities (nodes) are depicted as spheres (with radii proportional to community size) and the allosteric coupling between communities is represented by lines (edges) with a width corresponding to the degree of correlation. Red highlighted spheres indicate communities containing the FG-helix. Color-index and nomenclature of communities are indicated below the panel. **(B)**TYK2 amino-acids corresponding to R spine-, C spine-, R shell-forming residues²² and to residues serving dynamic bridging of the spines in active EPKs² along with bars colored according to their

corresponding communities identified in this study (positive only: $C_{ij} > 0.5$ and positive/negative: $|C_{ij}| > 0.5$, correlations-based community map analyses). Additional amino-acids with long hydrophobic side-chains that emerged from this study as potential contributors to dynamically supporting the TYK2 R- and C-spines in response to ATP.1Mg²⁺-binding and AL-phosphorylation, are also listed along with their corresponding communities.

Figure 5 . Side-chain node-betweenness centrality plot . Edge-betweenness centrality values for representative side-chain atoms (BC^{SC}), as obtained by Bio3D, are plotted along the TYK2-KD primary structure. Labels indicate amino-acids with higher BC^{SC} values ($>$ average $BC^{SC} = 356$). Secondary structure elements and functionally important regions and motifs discussed in the text, are boxed and labeled (as in Fig. S1). Asterisks and x's indicate TYK2 amino-acids mutated in cancer (missense or splice-site *Tyk2* mutations, respectively).

Figure 6. Community-based illustration of the allosteric communication network of amino-acids identified in the simulated TYK2-KD. Side-chains characterized by higher BC^{SC} values (as in Table I); namely, (A) significantly higher BC^{SC} values (> 1000 ; bold in Table I); i.e. with a pivotal role in mediating intra-KD allosteric signaling or (B) $BC^{SC} > 356$ (labelled in Fig.5), are mapped on the representative MD-snapshot and depicted as surfaces colored according to their corresponding SC community (as in Fig.4A). The main-chain of the simulated TYK2-KD is depicted as cartoon, colored in grey (in panel A) or according to the corresponding MC communities (in panel B). The ATP molecule and magnesium ion are shown in sticks and as a green-sphere, respectively. Schematic illustrations in panel B, indicate the position of other TYK2 sub-domains (pKD¹⁷; FERM-SH²¹⁶) in the autoinhibited full-length monomeric TYK2 (based on¹⁸). Encircled are amino-acids in the substrate binding region that are central to the identified communication network. Asterisks and x's as in Fig.5.

Figure 7. Part of the allosteric communication network of amino-acids mediating dynamic coupling between substrate- and ATP-binding sites in the simulated TYK2-KD. (A) Details of the residue interaction network, mapped on the 2-3 μ s representative conformation. Main-chain (in ribbon) and side-chains atoms (in sticks) are colored according to their corresponding communities (as in Fig.4A). Red- and blue-boxed labels indicate R- and C-spine forming residues, respectively. Dashed lines represent hydrogen bonding interactions. (B) Conformational ensemble of selected amino-acids from panel A, revealed by overlapping of representative MD-snapshots in the time-ranges of 2-2.5 μ s and 2.5-3 μ s of the MD simulation.

Figures

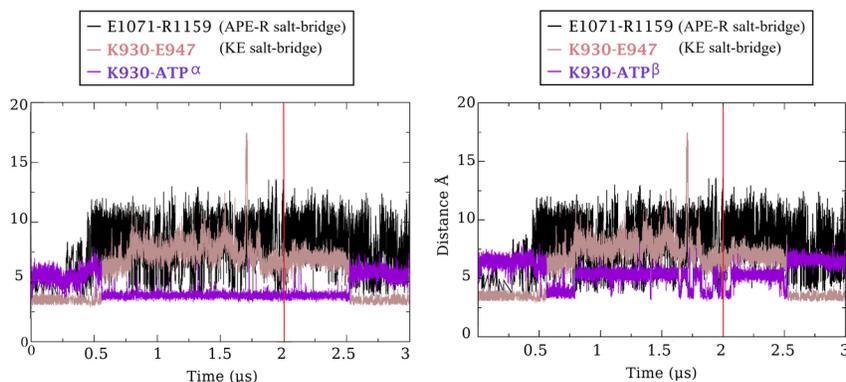


Figure 1.

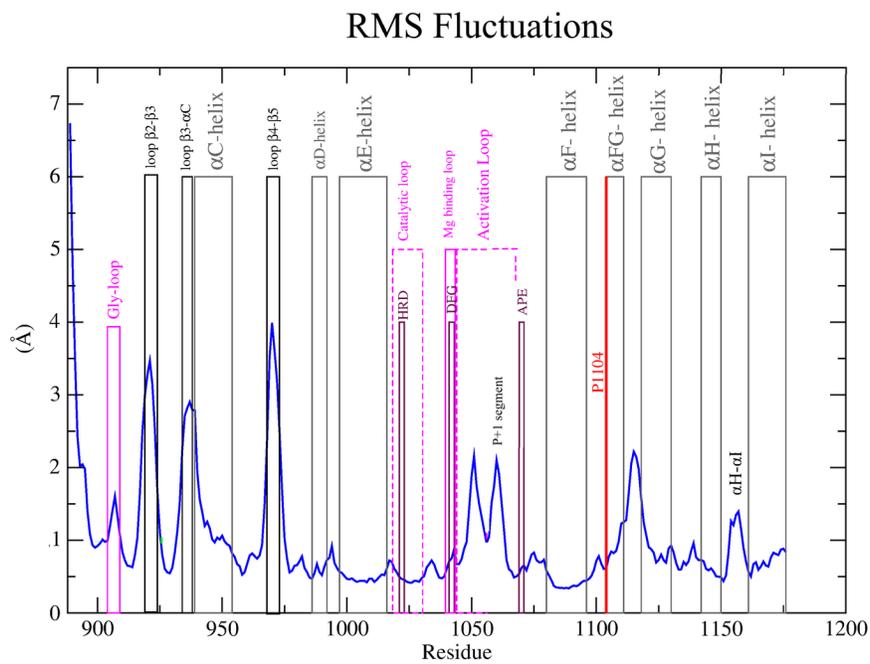


Figure 2.

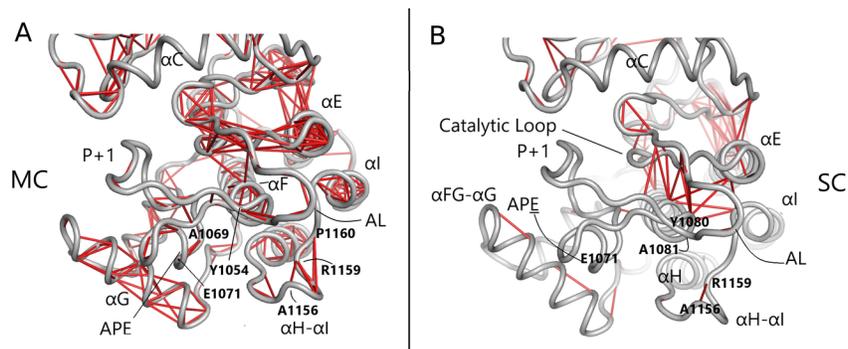


Figure 3.

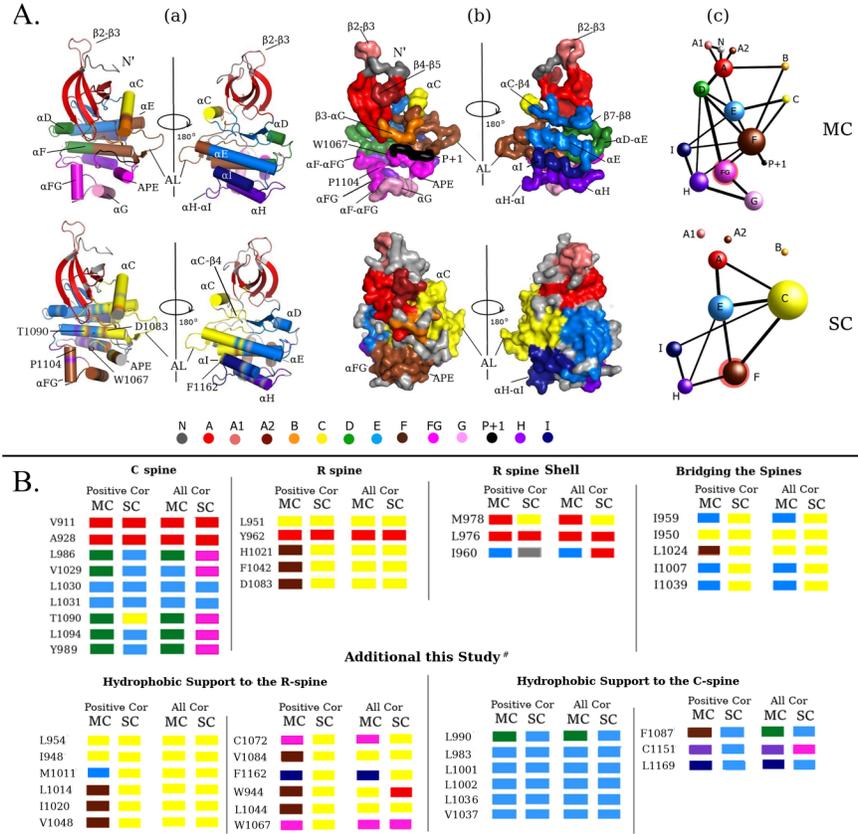


Figure 4.

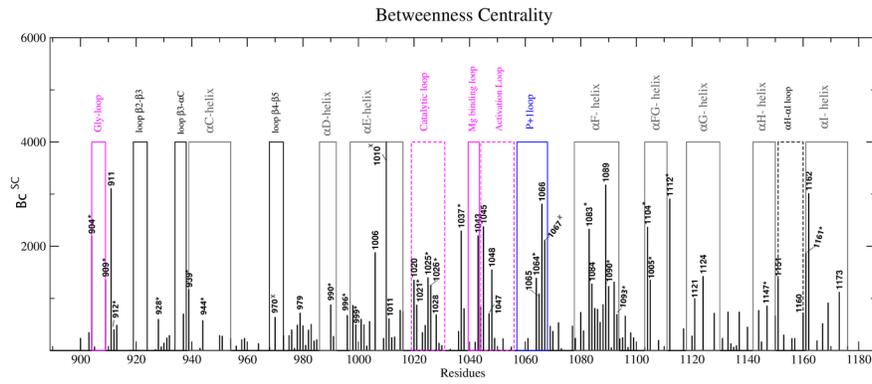


Figure 5.

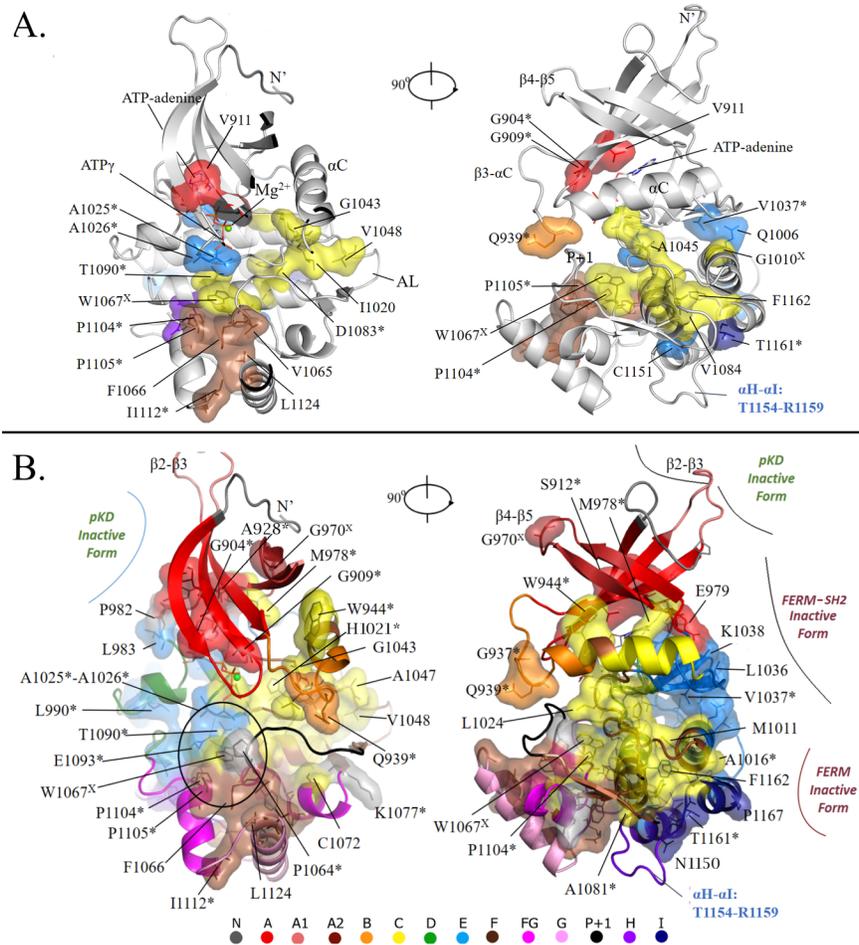


Figure 6.

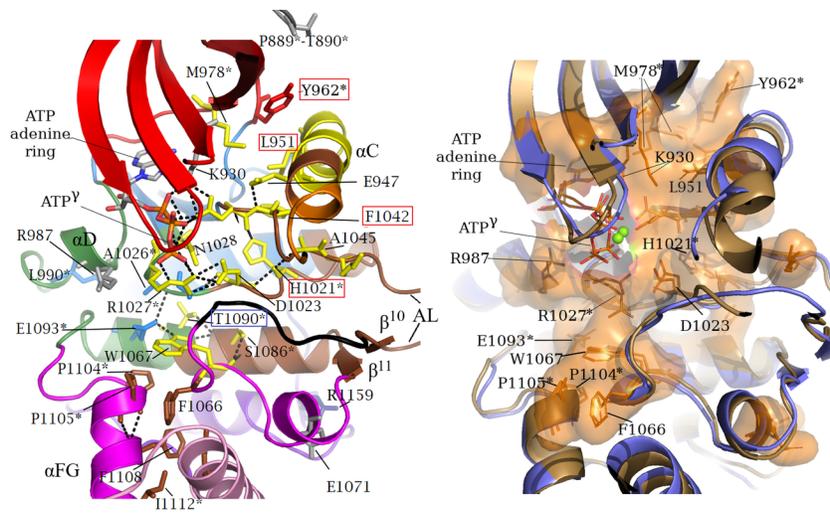


Figure 7.

Supporting Information

The file `Supplementary_Info_Lesgidou-Vlassi.docx` includes:

Supplementary Figures S1 to S9 with their captions and references

Conflict of interest

The authors declare no conflict of interest.