HHS Public Access

Author manuscript

Trends Biochem Sci. Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:

Trends Biochem Sci. 2015 November ; 40(11): 628–647. doi:10.1016/j.tibs.2015.09.002.

Dynamics driven allostery in protein kinases

Alexandr P. Kornev[1] and **Susan S. Taylor**[1,2]

[1]Department of Pharmacology, University of California at San Diego, La Jolla, CA, 92093, USA

[2]Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, CA, 92093, USA

Abstract

Protein kinases have very dynamic structures and their functionality strongly depends on their dynamic state. Active kinases reveal a dynamic pattern with residues clustering into semirigid communities that move in μs-ms timescale. Previously detected hydrophobic spines serve as connectors between communities. Communities do not follow the traditional subdomain structure of the kinase core or its secondary structure elements. Instead they are organized around main functional units. Integration of the communities depends on the assembly of the hydrophobic spine and phosphorylation of the activation loop. Single mutations can significantly disrupt the dynamic infrastructure and thereby interfere with long distance allosteric signaling that propagates throughout the whole molecule. Dynamics is proposed to be the underlying mechanism for allosteric regulation in protein kinases.

Keywords

Protein kinases; allostery; protein dynamics; community analysis

Allostery and dynamics

The term “allostery” was introduced more than 50 years ago to describe the regulation of enzyme activity by molecules that are sterically distinct from the substrate[1]. These molecules bind to a site different from the substrate binding site, termed the “allosteric site” (from the Greek “ἄλλος” for “other/different”). Since then, the term has been universally accepted and currently is used to describe any type of long distance signaling (known as “allosteric signaling”) inside proteins, related not only to small molecule binding but also to covalent modifications, protein-protein interactions or even mutations [2–4]. Despite significant progress in understanding of the underlying principles of allosteric signaling, it has to be studied on a case-by-case basis as the signal transmission mechanisms can vary. In general, such transmission can happen in two different ways: in a sequential set of events that propagates linearly from the allosteric to the active site, or by modification of a dynamic

Corresponding authors: Kornev A.P. (akornev@ucsd.edu); Taylor S.S. (staylor@ucsd.edu).

profile of the protein (Box 1). A characteristic feature of the both models is a requirement for significant structural changes of the enzyme. However, it was demonstrated subsequently that allostery can also be driven not by large motions of the protein, or by folding/unfolding of its domains[5] but by a change in the dynamic properties of residues with no significant conformational changes of the whole protein. The case of so called “Entropy driven allostery” predicted theoretically by Cooper and Dryden [6] was later confirmed for calmodulin [7], PDZ domains [8], the catabolite activator protein [9] and the Trp RNA-binding attenuation protein [10]. In all these cases protein dynamics is an essential key requirement for the allosteric regulation.

In this review we focus on the dynamic properties and allosteric regulation of protein kinases, which constitute nearly 2% of the human protein-coding genome [11] and are one of the most important protein families as they regulate much of biology. Protein kinase dysfunction is associated with multiple diseases including cancer, vascular disorders, and diabetes as well as neurological and inflammatory diseases. This makes protein kinases attractive therapeutic targets. Currently 28 small-molecule protein kinase inhibitors are approved by the U.S. Food and Drug Administration (FDA) , and thousands of clinical studies targeting kinases are ongoing [12]. Protein kinases modify about a third of all proteins in human cells [13] by transferring the γ-phosphate of ATP to the hydroxyl group of serines, threonines or tyrosines. Such phosphorylations result in significant changes of structural features and/or function of the proteins and are considered to be classic allosteric signaling events [1, 14, 15]. Protein kinases themselves can be regulated in multiple ways, that includes interaction with regulatory subunits[16], autoinhibition[17], dimerization[18], but the “classical” mechanism of their regulation is via phosphorylation either by other kinases or by autophosphorylation. In this review we will focus on protein kinase regulation limited to their kinase domain only. From the functional point of view protein kinases are classified as transferases but it is important to distinguish them from conventional metabolic enzymes. The latter have evolved to be very efficient catalysts with turnover numbers ranging from $\sim 40 s^{-1}$ for hexokinase IV [19] up to $530{,}000 s^{-1}$ for carbonic anhydrase [20]. Protein kinases do not operate as metabolic factories; instead, they serve as molecular regulators of these production lines switching them on and off. The most critical property for any switch is its ability to change its state quickly and unambiguously upon receiving a signal. Thus, protein kinases, being very poor transferases with typical turnover rates between $0.2 s^{-1}$ (e.g. for Src [21]) and $20 s^{-1}$ (e.g. for PKA [22]), have evolved to be highly malleable molecular switches that allow them to sense a wide range of signaling agents such as nucleotides, lipids, ions, sugars, hormones and growth factors. This “switch-like” character of protein kinases is mediated by their extraordinary structural plasticity and by their dynamic properties.

Internal architecture of protein kinases is built around dynamically assembled hydrophobic spines

All eukaryotic protein kinases have a structurally conserved core of approximately 250 residues that fold into a bilobal structure with a smaller N-terminal lobe (N-lobe) and a larger C-terminal lobe (C-lobe). A deep cleft between the lobes serves as a binding pocket

for the adenine ring of ATP (Figure 1). Comparison of protein kinase structures with and without ATP showed that the connection between the lobes is flexible. The N-lobe in the apo-form moves away from the C-lobe up to 9Å thereby opening the ATP binding cleft [23]. Such flexibility is considered to be important for nucleotide exchange during the catalytic cycle [24–27]. Flexibility is mediated by two ensembles of hydrophobic residues that connect the lobes by traversing the kinase core from the large αF-helix that spans the center of the C-lobe to the rigid β-sheet in the N-lobe [28–30]. They were termed "spines" to emphasize their flexible nature analogous to spines in vertebrates. One of the spines is called the Regulatory spine (R-spine) as it is assembled in all active kinases, and its disassembly leads to kinase inactivation. The second spine is termed the Catalytic spine (C-spine) as it is completed by the adenine ring of ATP which then poises the kinase for catalysis. The spines are a pair of unconventional structural elements that, although they are conserved in all EPKs and can even be traced in APKs [31], were not initially detected as they do not follow classical sequence motifs or secondary structural units. Their hydrophobic nature is the key property that allows for the kinase core mobility, a feature that, as we will show, is vital for protein kinase functionality. Here we revisit the conserved motifs of the kinase core paying special attention to key hydrophobic residues whose functional significance was not previously appreciated.

N-lobe

The N-lobe contains a five-stranded β-sheet with a large mobile αC-helix inserted between β3 and β4. The sheet carries a set of hydrophobic residues that accommodate the adenine ring of ATP. Two of them, Val57* from β2 and Ala70 from β3, belong to the C-spine and are conserved in EPKs [32]. The N-Lobe also contains two R-spine residues, Leu106 at the beginning of β4 and Leu95 at the end of αC.

G-Loop—The most dynamic loop in the kinase core, called "Glycine-rich Loop" or "G-Loop", lies between β1 and β2 and contains three highly conserved glycines (Gly50, Gly52 and Gly55) that form the GxGxxG-motif. The main function of this loop is to position the adenine ring and the phosphates of ATP so it is critically important. The G-Loop is associated with multiple disease mutations [33]. Sometimes it is incorrectly identified as a P-Loop, also known as the Walker motif, which is a recurring ATP phosphate binding motif in enzymes such as myosin and adenylate kinase [34]. However, the consensus sequence of the Walker motif is GxxxxGK(T/S) with the only similarity to the G-Loop being two glycines in positions 50 and 55 (Figure 2). The conserved Lys at the end of the P-Loop helps to position the γ-phosphate of ATP while in the EPKs this residue (Arg56 in PKA) reaches outwards to couple the G-Loop to other distal segments that typically contribute to closing of the G-Loop. The polar threonine/serine residue that follows the Lys in the P-Loop is almost universally a hydrophobic valine in the EPKs and is a part of the C-spine (Val57). In addition, the G-Loop always connects two β-strands, while the P-Loop connects a β-strand and an α-helix [35]. The entire P-Loop is committed to binding ATP and positioning the γ-phosphate for transfer at the base of the active site cleft, while in the EPKs the entire

*Here and further we use numbering for the prototypical cAMP-dependent kinase (PKA) unless otherwise specified

molecule contributes to positioning the γ-phosphate, which is at the edge of the active site cleft.

β3 Strand—β3 in the EPKs, ELKs and APKs contains a universally-conserved lysine (Lys72) [32] that coordinates the α- and β-phosphates of ATP and anchors the αC-helix via a salt bridge to a conserved glutamate (Glu91). This lysine is one of the most critically important residues. Its mutation even to a similar arginine usually leads to almost complete kinase inactivation [36]. In EPKs it is flanked by two hydrophobic residues: Ala70, which is a part of the C-spine and Leu74 [32]. The role of the latter was not appreciated until the recent discovery that it stabilizes the αC-helix. In Bruton's tyrosine kinase (Btk) substitution of this leucine (Leu432 in Btk) (Figure 3B) by isoleucine led to increased in activity [37]. Mutation of the analogous leucine in serine/threonine-protein kinase B-raf (BRAF) (Leu485 in BRAF) to phenylalanine is found in Cardio-facio-cutaneous syndrome [38], and replacing this Leu with Phe, Met or Ile in BRAF and serine/threonine-protein kinase C-raf (CRAF) leads to constitutive activation [39].

αC Helix—This prominent α-helix from the N-lobe is conserved in all EPKs, APKs and ALKs [32, 40]. Being a part of the N-lobe sequence, it is positioned between the two lobes and serves as an important structural integrator that defines the dynamic behavior of the whole molecule. It contains the RS3 residue from the R-spine and any displacement of the helix leads to disruption of the spine and inactivation of the kinase. Stability of the αC-helix is, therefore, of great importance and is attained by multiple contacts that tightly surround the helix both from the inside of the kinase core and from flanking regions that surround the core (Figure 3) [41]. The stabilizing role of the conserved Lys72/Glu91 salt bridge and Leu74 from the β3-strand was discussed earlier. Another β-strand from the N-lobe, β5, carries two residues: Met120 (known as "the gatekeeper") and Met118. They are a part of the so called "shell" that surrounds the R-spine, interact with the RS3 residue and are collectively important for kinase regulation [42]. The gatekeeper residue in serine/threonine kinases is predominantly a large aliphatic residue like methionine or leucine, whereas in majority of tyrosine kinases and tyrosine-like kinases it is a small polar threonine. Substitution of this threonine in tyrosine kinases by a large hydrophobic methionine or isoleucine leads to constitutive activation of the kinase and promotes malignancy [43]. Similarly, replacement of the RS3 Leu in BRAF/CRAF with Phe or Met leads to constitutive activation and creates an oncogene [39, 44]. The mobility of the spine residues and the residues flanking the spine are thus an essential element for dynamic regulation. The C terminus of the αC-Helix is anchored to the αC-β4 loop while the N terminus is often anchored through a basic residue to the phosphorylation site (Thr197) in the Activation Loop [30].

Residues that lie outside of the kinase core also stabilize the αC-helix. These can come from either their N- or C-terminal tails as in PKA and BRAF or from protein-protein interactions as in the cyclin-dependent kinases (CDKs) [14, 45]. Several pockets around the αC-helix are conserved in their size and position in different kinases [46]. A classic example is the "hydrophobic pocket" at its N terminus, discussed earlier, that can be filled in different ways. Kinase-specific examples are the "hydrophobic motif" (FxxF) from the C-terminal tail

of AGC kinases [47], cyclin in CDKs [48] and the “asymmetric dimer” in Epidermal growth factor receptor (EGFR) [49]. Small molecules can also be targeted to this pocket [50]. Another pocket lies at the C terminus of the αC-helix right above the αC-β4 Loop. Its regulatory role is well-studied in Src and Hck tyrosine kinases. In both cases this pocket interacts with a conserved tryptophan in the N-terminal linker. Interestingly, mutation of this tryptophan to alanine leads to opposite outcomes: in the case of Src it is an activating mutation, but in Hck it is inactivating [51]. As the pocket is positioned at the pivot point for the C-helix, it would be very difficult to initiate the movement of the whole bulky helix at this site. It is much more plausible to use this pocket for the helix stabilization by either inserting a “wedge” between the two R-spine residues (RS3 and RS4) in case of an inactive kinase, or, conversely, by reinforcing the R-spine of the active kinase via hydrophobic interaction with the tryptophan indole ring.

Another C-terminal pocket that can control the dynamics of the αC-helix lies on the opposite side of the helix in close proximity to the β9 strand (Figure 3). In PKA it is filled with the side chain of Trp30 from the N-terminal tail, and mutation of Trp30 to alanine or removal of the helix leads to destabilization of PKA [52]. Similar contacts can be observed in other kinases. In Aurora A kinase the analogous side chain of Phe35 is carried by an activating microtubule-associated protein TPX2 [53], in Casein kinase IIα Tyr23 comes from the N-terminal tail [54], and in MAP kinase ERK2 it is the C-terminal tail that delivers Phe327 to this pocket[55]. BRAF does not have a well-defined concave pocket at this position but interactions around the β9-αC site play an important regulatory role. The β9-strand carries a very important Val600. Although the V600E mutation is the most common mutation associated with melanoma[56], V600F also stabilizes the C-helix and leads to constitutive activation of BRAF[39] indicating once again the importance of hydrophobic substitutions that can shift the balance between active and inactive states and hijack a complex regulatory mechanism.

αC-β4 Loop—The αC-β4 Loop is another essential highly conserved loop that connects the αC-helix and β4 [32]. This loop is flanked by the two N-lobe R-spine residues (RS3 or Leu95 in the αC-helix and RS4 or Leu106 in the β4-strand) while the structurally conserved β-turn lies between Phe100 and Leu103. The αC-β4 Loop is tightly anchored to the αE-helix via an EPK-specific tyrosine/phenylalanine (Tyr156) and is the only part of the N-lobe that, from the dynamics point of view, belongs to the C-lobe [57]. It serves as a pivot point for the C terminus of the αC-helix and allows the helix to swing in and out during the activation/inactivation process (Figure 4). Analysis of the conserved interactions of this loop with the core of different EPKs suggested that it not only modulates ATP binding but can also play an important regulatory role [32]. It contains a conserved Val104 that was shown to be an important stabilizing residue for the R-spine [42]. In EGFR substitution of Arg776 (105 in PKA) to histidine, an αC-β4 mutation found in lung cancer, turns this kinase into a “superacceptor” for dimerization dependent activation[58]. In RAF kinases the αC-β4 Loop contains a critical arginine (position 99 in PKA) that is responsible for kinase activation via dimerization [59], whereas in ErbB2 two glycines from this loop (97 and 99 in PKA) define the activation state of this kinase [60] (Box 2).

C-lobe

While the N-lobe structural elements can be found in all EPKs, ELKs and even in most APKs, the C-lobe has a set of unique elements that are specific for EPKs. Out of six helices conserved in EPKs (αD through αI) only the first three (αD, αE and αF) are present in ELKs. Three helices αG, αH and αI (also known as GHI helical subdomain) together with the large loop region between the αE- and αF-helices are considered to be a product of later (prokaryotes vs. eukaryotes) evolution that turned ELKs into dynamic and allosterically regulated switches [45]. Four short β-strands form two β-sheets that stabilize the geometry of the extended loop region. The β6-β9 sheet is usually a signature of the active conformation and β9 is often disassembled in inactive kinases. The β7-β8 sheet is present in active and inactive EPKs as β7 is an essential part of the C-spine that is always assembled. β7 consists of three hydrophobic C-spine residues; the middle residue serves as an anchor site for binding the adenine ring of ATP (Leu173 in PKA). A short Catalytic Loop is stretched between the R- and the C-spines and contains most of the critical catalytic residues (Asp166, Lys168 and Asn171) (Figure 1G).

A long Activation Segment (~35–40 residues in most kinases) starts with the universally conserved Asp-Phe-Gly (DFG) motif between β8 and β9 and ends at the αF-helix (Figure 1). The DFG-aspartate is catalytically important as it binds a magnesium ion implicated in catalysis [25] and nucleotide exchange [27]. The DFG-motif is also called the “Magnesium Positioning Loop” [25]. The next part of the Activation Segment is called Activation Loop. This is the most diverse and flexible part of the whole kinase core. Often it undergoes an order/disorder transition. It typically contains a phosphorylation site that is the centerpiece of protein kinase allosteric regulation. After phosphorylation, it forms a set of strong hydrogen bonds to key elements of the kinase core: N terminus of the αC-helix, β9-strand, and the HRD-arginine from the Catalytic Loop. This leads to the correct assembly of the R-spine and proper orientation of the DFG-aspartate and correct alignment of the αC-Helix [30].

Following the Activation Loop is the so called “P+1 Loop” that is located in close proximity to the γ-phosphate of ATP. It serves as a docking site for the backbone of the substrate P-site residue and also for the side chain of the P+1 residue that follows the phosphorylation site*. The C-terminal portion of the P+1 Loop is stabilized by a conserved tyrosine in serine/threonine kinases (Tyr204) or tryptophan in tyrosine kinases that is tucked in between the αF- and αG-helices. An apparent role of the loop is to stabilize the substrate and correctly position the receiving hydroxyl group. One would thus expect that mutation of Tyr204 would influence only substrate binding, but its mutation to alanine in PKA leads to general destabilization of the molecule, a dramatic change of its catalytic activity and global dynamic changes registered by NMR [61–63]. As discussed later, the P+1 serves as a link between protein tethering to the αH-αI loop and the active site. Not surprisingly the P+1 Loop was found to be one of the “hotspots” for disease related mutations [33].

*The substrate residue to be phosphorated is known as the “P site”, hence, the flowing residue would be “P+1”

The highly conserved Ala-Pro-Glu (APE) motif follows the P+1 Loop. Its alanine and proline are anchored to a conserved tryptophan in the αF-helix (Trp222) while the APE-glutamate is connected via a buried salt bridge to a conserved arginine (Arg280) in the loop between the αH- and αI-helices. The bridge plays a significant structure-defining role as its mutation in PKA leads to destabilization of the kinase structure and a decrease of catalytic efficiency [64]. Mutation of the APE-motif residues as well as their binding partners (Trp222 and Arg280) were implicated in numerous diseases [33].

The EPK-specific αG-helix together with a portion of the Activation Segment right after the APE-motif serves as a docking site for other protein substrates [65] and regulatory proteins [66], as well as phosphatases [67] that dephosphorylate the Activation Loop and inhibit the kinase. Although the αH-and αI-helices are universally conserved in all EPKs, their role is often not clear, mainly because they are so distant from the active site of the kinase. The αI-helix of c-Abl tyrosine kinase was shown to be a mediator of myristyl-dependent activation [68]. Herman and coworkers showed that in the yeast homolog of PKA, Tpk1, the αH-αI loop serves as an allosteric tethering site for substrate proteins that allows for communication with the active site of the kinase [69]. The αH-helix of EGFRs is also thought to be an allosteric activator in the “asymmetric dimer” complex [49].

In conclusion, protein kinases have paradoxical structural properties. On the one hand they have two lobes that are very mobile with respect to each other. On the other hand, to perform their function, protein kinases have to attain a very specific configuration and to maintain it long enough for the phosphotransfer to occur. This contradiction is reconciled by unconventional structural ensembles known as hydrophobic “spines”. They provide a firm but, at the same time, flexible connection between the lobes, retaining their mobility but also positioning the γ-phosphate of ATP and the substrate peptide in a very precise way. Hydrophobicity of the connections inside the kinase core is, thus, a vital property that allows for both precision and flexibility. It is also important that these hydrophobic motifs are not contiguous as they need to be dynamically assembled to provide fast and reliable regulation of the kinase activity. Most importantly the nature of the mosaic of hydrophobic residues that flank the αC-Helix and the spines is not absolutely conserved; each kinase has a unique hydrophobic space that allows for a myriad of regulatory mechanisms. Replacing a small hydrophobic residue with a larger one such as Phe, Ile, or Met, however, can easily hijack the entire finely tuned regulatory mechanism and thereby create an oncogene. While the dynamic nature of a kinase molecule has clear importance for regulation, a set of elegant NMR studies has helped to shed light on its significant role in allostery and catalysis[26, 70–72].

Protein kinase dynamics portrayed by NMR

Although X-ray crystallography has been essential for teaching us about protein kinase structure and function, it has a substantial limitation in that it provides only static snapshots of a structure. In many cases crystallization of a mutant with significantly different functional properties yielded very similar structures compared to the wild type kinase. For example, ERK2 is activated by double phosphorylation of its Activation Loop, and this leads to a drastic increase in activity from $0.0002s^{-1}$ to $10s^{-1}$ [73]. However, comparison of the

active and inactive structures shows that the inactive kinase looks very similar to active ERK2, having a correctly assembled R-spine and only minor changes in the substrate binding site. Similarly, mutation of Tyr204 in PKA to alanine, as described in the previous section, leads to global changes in function, but the crystal structure of the mutant is virtually identical to the wild type kinase with the root-mean square deviation (RMSD) for over 300 C_α atoms being 0.4Å [74]. Long distance effects on flexibility of the αC-β4 Loop were reported for PKA [75]. This study measured anisotropy decay of fluorescein maleimide bound to the loop at the 99 position. Myristylation of the PKA N terminus significantly reduced flexibility of the site, but its flexibility increased significantly after ATP/PKI binding. Nevertheless, the geometry of the αC-β4 Loop and the overall structure of the kinase core did not reveal any significant changes related to the myristylation or the nucleotide binding, suggesting a complex dynamics based communication between these three sites. Another example is the allosteric inhibitors GNF-2 and GNF-5 that bind to the myristyl binding pocket of BCR-Abl kinase that is located about 20Å away from the active site. Although their binding does not lead to major structural changes in the kinase core, it nevertheless inhibits activity and changes the properties of the ATP binding site [76]. How can information about a posttranslational modification, tethering of a protein substrate, or drug binding be sensed globally throughout the kinase molecule without significant conformational changes? How also can a single oncogenic mutation interfere with this network of communication? Is it possible that the “violin model” is employed (Box 1) and the long distance signaling is based on changes in the dynamic portrait of the whole molecule? To answer this question one has to evaluate dynamic properties of the protein kinase before and after the modification. This task can be tackled by NMR, a powerful technique that can provide insight into protein dynamics and help to understand long distance allosteric effects.

For a long time implementation of NMR for protein kinases was limited due to the large size of these enzymes, but recent advances in backbone amide assignments [77] as well as introduction of new techniques such as NMR relaxation dispersion methods [78], allow us to take a closer look at the protein kinase dynamics. An extensive set of NMR studies, performed on PKA [26, 70–72, 79], demonstrated how closely dynamics is engaged in the catalytic function of PKA. It was known early on that ATP binding leads to closure of the active site and makes the kinase core more stable [80, 81]. It would thus be logical to suggest that such closing would cause stabilization of the molecule and, hence, a decrease in dynamics. Surprisingly, nucleotide binding caused a significant increase of slow conformational exchange in the μs-ms timescale. Furthermore, binding of the phospholamban peptide substrate did not eliminate the dynamics. On the contrary, the slow dynamics of the substrate itself went up. This was consistent with isothermal titration calorimetry that indicated that entropy is the driving force for the phospholamban binding to PKA [26]. Florescence anisotropy studies of PKA also linked entropy to phosphotransfer [82]. This is also in line with the recent discoveries that highlight the leading role played by entropy in calmodulin substrate recognition, inhibitor binding to lysozyme and dihydrofolate reductase catalysis [83].

For PKA the dynamic changes were not localized around the ATP or substrate binding site. Instead, the slow conformational exchange was registered throughout the protein: the N-

lobe, Activation Segment and C-terminal tail (Figure 5A-5C). NMR analysis of PKA in different states of "compactness", from the most open, flexible apo form, to the most compact form with inhibitor peptide, nucleotide and two metal ions bound, showed that the difference between the square of the chemical shifts ($\Delta\omega^2$) for many residues is linearly correlated to the rate of their slow relaxation rate of conformational exchange (R_{ex}). This allowed the authors to suggest that the emergence of the slow dynamics that occurs following nucleotide binding is related to the closing and opening of the kinase lobes. According to the proposed model, assembly of both hydrophobic spines prompts a transition of the kinase into a new dynamic state, termed "dynamically committed" versus the "dynamically uncommitted" state observed in the apo form. It is important to distinguish this terminology from the traditional "active" vs. "inactive" options as the dynamically uncommitted kinase has all the structural features of an active kinase including phosphorylation of the Activation Loop, an assembled R-spine, and an unobstructed substrate binding site, and thus cannot be called "inactive" in a conventional way. Completion of the C-spine by the adenine ring of ATP mediates this global transition.

Although the direct involvement of slow dynamics in catalysis remains to be proven experimentally, several observations strongly support this hypothesis. First, the μs-ms frequency of the emerging motions in the kinase is very important as it is related to essential biological processes such as folding, large domain motions, and catalysis. Indeed, according to the estimations made by Veglia and coauthors, the average conformational exchange rate in PKA ($\sim$20–30s^{-1}) is very close to the k_{cat} of PKA ($\sim$23s^{-1}) [26]. Earlier kinetic studies showed that slow conformational changes partially control the turnover rate of PKA [24], emphasizing the importance of dynamics in catalysis. It is, thus, logical to suggest that the slow dynamics observed in the NMR experiments and in the kinetic studies are related to the same phenomenon. In addition, NMR studies of the Tyr204Ala mutant of PKA showed a significant disruption of the slow dynamics throughout the molecule [72]. As pointed out earlier, this mutation leads to a significant reduction of catalytic rate without substantial changes in the structure of the C:PKI:ATP. Apparently, the explanation for the low kinase activity in this case lies not in the realm of structural changes but can instead be attributed to disruption of the dynamic properties of the kinase. Finally, a very similar phenomenon to the "dynamically committed" state of PKA was observed in ERK2 [84] when the dynamic properties of the methyl groups of leucines, isoleucines and valines were compared in active and inactive ERK2. By measuring relaxation dispersion in the μs-ms timescale they showed that the kinase core of inactive ERK2 has very diverse dynamics with relaxation times ranging between 900 and 2200 s^{-1}. Activation of the kinase by phosphorylation of the activation loop led to a drastic change in the dynamics. Nineteen out of twenty two residues, distributed all around the kinase, showed slow conformational exchange with rates $\sim$300 s^{-1} (Figure 5D-5E). In this case, the transition from an inactive to an active conformation leads to the emergence of slow and global dynamics of the kinase core.

Insights presented by NMR confirm the critical importance of dynamics for protein kinase function. In many cases, static snapshots of kinase structures obtained from X-ray crystallography fail to reflect functionally important changes caused by posttranslational modifications, mutations, ligand or protein binding. Although NMR can provide objective information on several important aspects of protein kinase dynamics, it cannot deliver the

complete picture of physical processes that drive kinase function and, thus, cannot explain the observed phenomenon of dynamic transitions that regulate protein kinase activity. Increasingly, due to major advances in technology, computational methods can successfully complement NMR experiments and allow us to analyze protein dynamics on an atomic scale. As we move forward it is becoming evident that both backbone and sidechain dynamics will be essential for comprehensive portraits of kinase dynamics.

Viewing protein kinase dynamics through computational lenses

Coarse-grained approaches

Several computational methods were developed for analyzing protein dynamics. The most simple approach called the Gaussian Network Model (GNM) or Elastic Network Model (ENM) represents a protein as a "ball-and-spring" network with Cα atoms connected to their neighbors by harmonic springs[85]. Despite a significant simplification of the system, this model, in combination with Normal Mode Analysis (NMA), provides us with a glimpse into very important features of protein long scale dynamic behavior [86]. Shudler and Niv used this approach to analyze the dynamic properties of four protein kinases in their active and inactive forms: two serine/threonine kinases (RAC-alpha serine/threonine-protein kinase (PKB) and Cyclin-dependent kinase 2 (CDK2)) and two tyrosine kinase (Tyrosine-protein kinase ABL1 (ABL1) and Insulin receptor tyrosine kinase (IRK)) [87]. They found that the kinase core can be partitioned into semi-rigid bodies, termed "blocks", of about five to eight residues each. The number of blocks (32 to 45) and their content varied depending on the kinase and their activation state. However, some blocks were consistently found in different kinases. The "pivot" block, which included the tip of the αC-β4 Loop (residues 99–105 in PKA), was detected in seven out of eight cases with the exception of inactive Abl. Three out of four active kinase structures revealed a "loop" block that includes the DFG-region of the Activation Segment and a part of the αC-helix (residues 88–96 and 184–188). This block contains the middle part of the R-spine (RS2 and RS3) and reflects stabilization of the assembled R-spine in an active kinase. However, no other blocks were found to be conserved between the kinases or had clearly identifiable function. This may be due to the simplified character of the model, because it did not include dynamics of the side chains.

Graph-theory based methods

Another approach for analysis of protein dynamics uses graph theory to represent protein molecules. In this case vertices residues whereas edges reflect interactions between them. This showed that such graph theoretical concepts as centrality, betweenness, and closeness can be related to important properties of a protein molecule and can identify regions that are critical for folding, stability and function [88]. Edges in the graphs can be defined in several different ways. An edge can be created if the residues are "in contact"; that is, they are in relatively close proximity in a particular protein structure. Molecular dynamics can be used to make the "contact" definition more reliable; that is, the contact is formed if the residues are close to each other during a reasonably long period during the Molecular Dynamics (MD) simulation. Cross-correlation matrix [89] or mutual information matrix [90] derived from MD simulations can also be used to define the edges. In these cases even more detailed information on protein dynamics can be included in the analysis.

Using interaction energy [91] as a criterion for the edge creation James and Verkhivker [92] analyzed 50 ns MD simulations of different forms of EGFR and ErbB4. They performed community analysis of these data based on the Clique Percolation Method [93] and detected a varying number (from 8 to 30) of semi-rigid bodies that consisted of four to eight residues. The number and the structure of these communities differed significantly between kinases and their functional states; however, generally larger and more stable communities are observed in the kinase core surrounding the R-spine and the αC-helix in active kinases, supporting observations made by Shudler and Niv [87].

Girvan-Newman Community analysis

Another study [94] used a community analysis based on the Girvan-Newman algorithm [95] and Mutual Information analysis [96] to study multiple microsecond scale simulations of PKA. Application of the Girvan-Newman algorithm is advantageous for allosteric signaling as it is based on the concept of edge betweenness, providing a quantitative estimate of allosteric coupling between communities. As shown earlier, changes of such coupling play an important role in allosteric regulation of *Staphylococcus aureus* Sortase A [90]. Another important feature of these studies compared to the earlier studies was that both amino acid main-chains and side chains were included in the analysis as separate entities, providing a more detailed picture of the kinase dynamics. . As emphasized earlier, one cannot exclude side chain dynamics.

Several large (40 to 60 residues) closely interconnected communities were detected in PKA bound to ATP and two magnesium ions (Figure 6). These major communities have clearly identifiable functions, which led to a suggestion that the organization of dynamic communities will be a universal feature for all protein kinases. Community A, for example, is in the N-lobe and is responsible for binding the adenine ring of ATP (Figure 7A). A portion of the seemingly rigid β-sheet of the N-lobe is split from the A community and constitutes a part of a relatively small but tightly interconnected Community B (Figure 7B). The function of Community B is to help position the αC-helix and the γ-phosphate of ATP for transfer. It is built around the “hydrophobic pocket” (Figure 4A) that plays a critical role in activation for many kinases. It is important to emphasize that Communities A and B include N- and C-terminal tail segments as well as the bulk of the N-lobe. Because these critical communities include both core and non-core residues, it is mandatory that they will differ for each kinase; while the conserved function will be maintained, the way in which those functions are regulated will be kinase-specific.

The αC-helix is at the center of Community C, which also is includes the DFG-motif, the β6-β9 sheet, and a significant part of the αE-helix. In PKA, Community C, also includes the C-terminal part of the αA-helix that fills one of the conserved hydrophobic pockets at its C-terminal end of the αC-helix (Figure 7C). This community forms strong communication bonds to the five largest communities and represent a central regulatory “hub” that is in a good agreement with an earlier assessment of the αC-helix as a “Signal integration motif” [25]. Both the “pivot” and the “loop” blocks detected by NMA [87] were included in Community C. Although most of the main-chains of the conserved and regulatory residues in the αC-β4 Loop were included in Community C, its side-chains were associated with the

neighboring communities. The side-chain of Leu106 that is a part of the R-spine (RS4) was a member of Community A, the side-chain of Val104 that is a part of the "Shell" (Sh1) was associated with the catalytic Community D, and for Phe102 at the turn in the αC-α4 loop both the main-chain and the side-chain belonged to Community E. This unusual property of the αC-β4 Loop reinforces its role as an important "communication hub" that mediates allosteric signals between the R-spine, the ATP-binding pocket and the C-lobe. Although the backbone of this loop is remarkably stable in all the conformational states, due in part to structured water molecules, it can sense most distal parts of the protein. Introducing a fluorescent probe at Asn99 demonstrated how sensitive this site is to ATP binding and to acylation [75]. This can explain the highly conserved structural features of this loop [32] and its involvement in regulating the activity of so many different protein kinases even though the sequence is not conserved [59, 60, 97].

The function of Community D is to properly position the ATP molecule, in particular the γ-phosphate and the two magnesium ions, so that the phosphate is poised for transfer to a protein substrate (Figure 7E,F). The importance of the second metal is essential for optimal transfer of the phosphate and its release is essential for releasing the ADP product [27]. Community D also includes most of the C-spine residues in the C-Lobe that are localized primarily in β7-strand and the αD-helix. Leu173 in β7 directly packs beneath the adenine ring while Ala70 and Val57 in Community A stack on top of the ring. Community D also includes the magnesium coordinating residues Asn171 in the Catalytic Loop and Asp184 in the DFG-motif drawing attention again to the importance of the metal ions. Tyr164 of the HRD motif is the only R-spine residue (RS1) that is part of Community D. Although its main chain is in Community C, its side chain is part of Community D. This residue, which bridges the R-spine to the αF-Helix through Asp220, is typically a histidine in most protein kinases. All of the backbones of the R-spine residues are in Community C as are the side chains of RS2 (DFG) and RS3 (αC) whereas the side chain of RS4 (β4) is in Community A. Community D through the HRD tyrosine thus integrates many communities and focuses them on the second magnesium ion.

Community F is the dominating community in the C-lobe and includes most of the EPK-specific features of the protein kinase superfamily. It is assembled around the Activation Segment with the Activation Loop phosphate and the conserved Glu208-Arg280 salt bridge as main structural organizers (Figure 7G), which are the defining features that distinguish the EPKs from the simpler eukaryotic-like kinases. This region defines the EPKs as molecular switches as opposed to efficient catalysts. The order-to-disorder transition of the Activation Loop is typically controlled by phosphorylation. Another highly conserved residue in this region is Tyr215, which positions the backbone of pThr197. This strategically positioned tyrosine, like Tyr156 in the αE-helix, appears to be essential. Many aspects of substrate peptide recognition at the active site are embedded within this community. The P +1 Loop is well recognized, however, Tyr204 in this loop reaches across and contributes to recognition of the P-2 residue. In addition, Gly200 and Thr201 interact directly with the P-site.

The C lobe also contains several minor communities (E, F1, G and H) that are known to be associated with substrate and/or inhibitor binding to peripheral sites, and the community

analysis shows clearly that they are also extremely sensitive to ligand binding in the ATP binding pocket. For example, removal of a single magnesium ion from the active site led to a significant re-organization of all the communities. Parts of the Communities H and E as well as the whole Community F1 and G merged with the Community F causing a considerable solidification of the C-lobe. The tails that flank the N-lobe also sense this metal ion. Clearly, the single metal ion does not lead to a global change in conformation yet the dynamic profile of the entire molecule is altered. It is likely that similar community-based signaling between the distant C-lobe areas and the active site underlies the allosteric effect observed in PKA [69]. Also consistent with this is mutation of the conserved Tyr204 that lies at the border of Community F and leads to a split of this community into two (F and F1) [72]. This, in turn, causes significant rearrangement of the key regulatory Community C with the middle of the R-spine separating into a separate Community C1. This disruption of the dynamic structure is reflected in the decrease in catalytic efficiency of this mutant [72].

Thus, the most recent computational results demonstrate that dynamics in protein kinases is precisely controlled. Residues do not move in a randomized and uniform way, instead, they form transient clusters (communities) that behave as semi-rigid bodies. Their size and the connectivity between them depend on multiple factors: the functional state of the kinase, bound ligands or even single ions, and tethered proteins. Changes of the dynamic profile can occur without significant changes of the kinase structure, but can have a profound effect on its function. These findings can explain long distance allosteric effects that could not be explained previously due to an absence of structural changes. The Community analysis also can explain many disease mutations that are distal from the active site.

Concluding remarks

All eukaryote protein kinases share a conserved core that contains twelve subdomains that were originally defined by Hanks and Hunter, and embedded within each subdomain are conserved sequence motifs [98]. Later, when the first protein kinase structure was solved [99], these motifs could be mapped onto the protein kinase secondary structures: α-helices, β-strands and loops [99], and those motifs correlated very well with the overall tertiary structure. Subdomain I, for example, correlates with β1 and β2 and includes the G-loop whereas Subdomain II corresponds with β3. The structure also allowed us to assign tentative functional roles to different parts of the molecule and to specific residues. In general, however, the role of hydrophobic side-chains was largely unrecognized and unappreciated. The discovery of the hydrophobic spine architecture introduced a new concept into how the kinase core is organized and how entropy might drive both allostery and catalysis. There is growing evidence that protein kinase function is mediated by the dynamic properties of its core, and the “hydrophobic spines” are a prerequisite for acquiring the dynamic mode termed “dynamically committed state” [26, 70] that is characterized by slow motions in the millisecond timescale that are uniformly spread throughout the whole protein kinase core. Analysis of protein kinase MD simulations shows that residues can form rather large semi-rigid blocks or communities, and this provides us with a new way to understand correlated motions and function that goes beyond simple sequence motifs and secondary structure. The G-Loop, for example, is now divided into two distinct functional units that integrate with other parts of the molecule. These communities undergo significant changes and

rearrangements as the kinase toggles through different functional states. Dissection of the kinase core into such functionally integrated communities can provide a new nomenclature that can be instrumental for analyzing protein kinase dynamics and, hence, their function. They can also be used as structural units for modeling allosteric signaling in parallel with the traditional quaternary, tertiary, secondary and primary elements of protein structure. Docking to one part of the community can easily be sensed, for example, by distal parts of the community. As the dynamics-based nomenclature reflects functional state of the kinase it can be applied in drug design as an in silico measure of a drug interaction with the target kinase. Disease mutations also validate the community analysis as these can either interfere with integration of the community or can drive the assembly of the active community and thus override the carefully choreographed regulatory mechanism. The new dynamics based nomenclature can also be used as a criterion

The community analysis also brings new understanding to classical symmetric allosteric complexes of oligomers [100], as well as symmetric and asymmetric dimers that are being discovered in so many protein kinases [44, 49, 59]. In addition, the community map analysis can define new ways for understanding how the conserved structural units that are embedded within the core of all kinases [101] are integrated with the flanking tails and domains to create a precisely controlled molecular switch. While the details of each community will differ somewhat with each kinase, the way in which the entire molecule is integrated in a concerted way with the highly conserved core machinery necessary to mediate the transfer of a phosphate will be conserved. We know that the internal architecture of protein kinases built around the hydrophobic spines is conserved. Now community analysis has revealed how the spines connect the functionally important communities providing lines of communication between them (Figure 8**, Key Figure**). With the dynamics studies we now have a window on these conversations. It is the differences that make each kinase unique and that make it challenging to identify allosteric inhibitors. The degree of conservation of the communities between different kinases is still an open question but it can be an important step towards better understanding of this important family of regulatory enzymes.

While a central dogma of biology is that structure will define function, we see here that there is an essential dynamic component that lies between structure and function. In some cases changes in function will correlate with changes in conformation such as ordering of the activation loop whereas in other cases when one considers the precise function of the active kinase there are no major conformational changes even though there are major changes in dynamic properties. Recent advances in NMR and particularly in computation are merging computational and experimental tools so that computational predictions can be experimentally validated. A whole world of dynamics lies between structure and function, and we are only just beginning to explore this world.

Acknowledgements

The authors would like to acknowledge support from the National Institutes of Health (GM19301 and GM100310) and from the Howard Hughes Medical Institute.

Glossary

Complex network (graph theory)	A large graph with non-trivial topological features where vertices of the graph represent a set of objects and edges represent connections between them.
Community (graph theory)	A set of vertices in a complex network that can be grouped into a subset of the network as the vertices in the community have more connections to each other than to the rest of the network
Community analysis	A computational method that partitions a complex networks into communities. The result is as a weighted graph that has a more simple structure than the original network and, thus, facilitates analysis of the latter
Edge betweenness	A number of the shortest paths between all possible pairs of vertices in a graph that pass through the edge
Girvan-Newman algorithm	One of several popular algorithms used to perform Community analysis. It uses the notion of edge betweenness to partition the complex network into communities. Its primary suggestion is based on the intuitively evident concept that if two communities are connected by a single edge, this edge will have the maximum edge betweenness because all of the shortest inter-community paths will have to go through this edge
EPKs	Eukaryotic protein kinases. A family of homologous serine-threonine and tyrosine protein kinases that contain 12 conserved sequence subdomains and have highly conserved catalytic core 3-dimentional structure
ELKs	Eukaryotic-like kinases. A family of phosphotransferases from prokaryotes that typically phosphorylate small metabolites and have sequence and structure similarities to the EPKs. They are considered to be evolutionary precursors of EPKs and have more diverse structures in comparison with EPKs
APKs	Atypical protein kinases. A relatively small family of eukaryotic phosphotransferases that share structural homology with EPKs and ELKs but are significantly different in sequence and structure. Their similarity to EPKs and ELKs is mostly related to the ATP binding features. They are considered to be an intermediate protein fold between protein kinases and other ATP binding proteins

References

1. Monod J, et al. Allosteric proteins and cellular control systems. J Mol Biol. 1963; 6:306–329. [PubMed: 13936070]
2. Gunasekaran K, et al. Is allostery an intrinsic property of all dynamic proteins? Proteins. 2004; 57:433–443. [PubMed: 15382234]

3. Nussinov R, Tsai CJ. Allostery in disease and in drug discovery. Cell. 2013; 153:293–305. [PubMed: 23582321]
4. Lu S, et al. Harnessing allostery: a novel approach to drug discovery. Med Res Rev. 2014; 34:1242–1285. [PubMed: 24827416]
5. Motlagh HN, et al. Interplay between allostery and intrinsic disorder in an ensemble. Biochem Soc Trans. 2012; 40:975–980. [PubMed: 22988850]
6. Cooper A, Dryden DT. Allostery without conformational change. A plausible model. Eur Biophys J. 1984; 11:103–109. [PubMed: 6544679]
7. Frederick KK, et al. Conformational entropy in molecular recognition by proteins. Nature. 2007; 448:325–329. [PubMed: 17637663]
8. Petit CM, et al. Hidden dynamic allostery in a PDZ domain. Proc Natl Acad Sci U S A. 2009; 106:18249–18254. [PubMed: 19828436]
9. Popovych N, et al. Dynamically driven protein allostery. Nat Struct Mol Biol. 2006; 13:831–838. [PubMed: 16906160]
10. McElroy C, et al. TROSY-NMR studies of the 91kDa TRAP protein reveal allosteric control of a gene regulatory protein by ligand-altered flexibility. J Mol Biol. 2002; 323:463–473. [PubMed: 12381302]
11. Manning G, et al. The protein kinase complement of the human genome. Science. 2002; 298:1912–1934. [PubMed: 12471243]
12. Wu P, et al. FDA-approved small-molecule kinase inhibitors. Trends Pharmacol Sci. 2015; 36:422–439. [PubMed: 25975227]
13. Cohen P. The origins of protein phosphorylation. Nat Cell Biol. 2002; 4:E127–E130. [PubMed: 11988757]
14. Johnson LN, Lewis RJ. Structural basis for control by phosphorylation. Chemical Reviews. 2001; 101:2209–2242. [PubMed: 11749371]
15. Kuriyan J, Eisenberg D. The origin of protein interactions and allostery in colocalization. Nature. 2007; 450:983–990. [PubMed: 18075577]
16. Taylor SS, et al. Assembly of allosteric macromolecular switches: lessons from PKA. Nat Rev Mol Cell Bio. 2012; 13:646–658. [PubMed: 22992589]
17. Antal CE, et al. Intramolecular C2 Domain-Mediated Autoinhibition of Protein Kinase C betaII. Cell Rep. 2015
18. Jura N, et al. Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. Mol Cell. 2011; 42:9–22. [PubMed: 21474065]
19. Heredia VV, et al. Glucose-induced conformational changes in glucokinase mediate allosteric regulation: transient kinetic analysis. Biochemistry. 2006; 45:7553–7562. [PubMed: 16768451]
20. Donaldson TL, Quinn JA. Kinetic constants determined from membrane transport measurements: carbonic anhydrase activity at high concentrations. Proc Natl Acad Sci U S A. 1974; 71:4995–4999. [PubMed: 4216027]
21. Lieser SA, et al. Phosphoryl transfer step in the C-terminal Src kinase controls Src recognition. J Biol Chem. 2005; 280:7769–7776. [PubMed: 15623523]
22. Grant BD, et al. Kinetic analyses of mutations in the glycine-rich loop of cAMP-dependent protein kinase. Biochemistry. 1998; 37:7708–7715. [PubMed: 9601030]
23. Akamine P, et al. Dynamic features of cAMP-dependent protein kinase revealed by apoenzyme crystal structure. J Mol Biol. 2003; 327:159–171. [PubMed: 12614615]
24. Adams JA. Kinetic and catalytic mechanisms of protein kinases. Chem Rev. 2001; 101:2271–2290. [PubMed: 11749373]
25. Johnson DA, et al. Dynamics of cAMP-dependent protein kinase. Chem Rev. 2001; 101:2243–2270. [PubMed: 11749372]
26. Masterson LR, et al. Dynamics connect substrate recognition to catalysis in protein kinase A. Nat Chem Biol. 2010; 6:821–828. [PubMed: 20890288]
27. Bao ZQ, et al. Briefly bound to activate: transient binding of a second catalytic magnesium activates the structure and dynamics of CDK2 kinase for catalysis. Structure. 2011; 19:675–690. [PubMed: 21565702]

28. Kornev AP, et al. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. P Natl Acad Sci USA. 2006; 103:17783–17788.
29. Kornev AP, et al. A helix scaffold for the assembly of active protein kinases. Proc Natl Acad Sci U S A. 2008; 105:14377–14382. [PubMed: 18787129]
30. Taylor SS, Kornev AP. Protein kinases: evolution of dynamic regulatory proteins. Trends Biochem Sci. 2011; 36:65–77. [PubMed: 20971646]
31. Vadas O, et al. Structural basis for activation and inhibition of class I phosphoinositide 3-kinases. Sci Signal. 2011; 4
32. Kannan N, Neuwald AF. Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? J Mol Biol. 2005; 351:956–972. [PubMed: 16051269]
33. Torkamani A, et al. Congenital disease SNPs target lineage specific structural elements in protein kinases. Proc Natl Acad Sci U S A. 2008; 105:9011–9016. [PubMed: 18579784]
34. Walker JE, et al. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1982; 1:945–951. [PubMed: 6329717]
35. Ramakrishnan C, et al. A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins. Protein Eng. 2002; 15:783–798. [PubMed: 12468712]
36. Iyer GH, et al. Consequences of lysine 72 mutation on the phosphorylation and activation state of cAMP-dependent kinase. J Biol Chem. 2005; 280:8800–8807. [PubMed: 15618230]
37. Boyken SE, et al. A conserved isoleucine maintains the inactive state of Bruton's tyrosine kinase. J Mol Biol. 2014; 426:3656–3669. [PubMed: 25193673]
38. Rodriguez-Viciana P, et al. Germline mutations in genes within the MAPK pathway cause cardio-facio-cutaneous syndrome. Science. 2006; 311:1287–1290. [PubMed: 16439621]
39. Hu J, et al. Kinase regulation by hydrophobic spine assembly in cancer. Mol Cell Biol. 2015; 35:264–276. [PubMed: 25348715]
40. Scheeff ED, Bourne PE. Structural evolution of the protein kinase-like superfamily. PLoS Comput Biol. 2005; 1:e49. [PubMed: 16244704]
41. Taylor SS, et al. PKA: a portrait of protein kinase dynamics. Biochim Biophys Acta. 2004; 1697:259–269. [PubMed: 15023366]
42. Meharena HS, et al. Deciphering the structural basis of eukaryotic protein kinase regulation. PLoS Biol. 2013; 11:e1001680. [PubMed: 24143133]
43. Azam M, et al. Activation of tyrosine kinases by mutation of the gatekeeper threonine. Nat Struct Mol Biol. 2008; 15:1109–1118. [PubMed: 18794843]
44. Hu J, et al. Allosteric Activation of Functionally Asymmetric RAF Kinase Dimers. Cell. 2013; 154:1036–1046. [PubMed: 23993095]
45. Taylor SS, et al. Evolution of the eukaryotic protein kinases as dynamic molecular switches. Philos T R Soc B. 2012; 367:2517–2528.
46. Thompson EE, et al. Comparative surface geometry of the protein kinase family. Protein Sci. 2009; 18:2016–2026. [PubMed: 19610074]
47. Kannan N, et al. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. Proc Natl Acad Sci U S A. 2007; 104:1272–1277. [PubMed: 17227859]
48. Jeffrey PD, et al. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature. 1995; 376:313–320. [PubMed: 7630397]
49. Zhang X, et al. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell. 2006; 125:1137–1149. [PubMed: 16777603]
50. Hindie V, et al. Structure and allosteric effects of low-molecular-weight activators on the protein kinase PDK1. Nat Chem Biol. 2009; 5:758–764. [PubMed: 19718043]
51. McPherson RA, et al. A different function for a critical tryptophan in c-Raf and Hck. Oncogene. 2000; 19:3616–3622. [PubMed: 10951567]
52. Herberg FW, et al. Importance of the A-helix of the catalytic subunit of cAMP-dependent protein kinase for stability and for orienting subdomains at the cleft interface. Protein Sci. 1997; 6:569–579. [PubMed: 9070439]

53. Bayliss R, et al. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. Mol Cell. 2003; 12:851–862. [PubMed: 14580337]
54. Niefind K, et al. Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme. EMBO J. 2001; 20:5320–5331. [PubMed: 11574463]
55. Canagarajah BJ, et al. Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. Cell. 1997; 90:859–869. [PubMed: 9298898]
56. Sun C, et al. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. Nature. 2014; 508:118–122. [PubMed: 24670642]
57. Tsigelny I, et al. 600 ps molecular dynamics reveals stable substructures and flexible hinge points in cAMP dependent protein kinase. Biopolymers. 1999; 50:513–524. [PubMed: 10479734]
58. Ruan Z, Kannan N. Mechanistic Insights into R776H Mediated Activation of Epidermal Growth Factor Receptor Kinase. Biochemistry. 2015; 54:4216–4225. [PubMed: 26101090]
59. Rajakulendran T, et al. A dimerization-dependent mechanism drives RAF catalytic activation. Nature. 2009; 461:542–545. [PubMed: 19727074]
60. Fan YX, et al. Mutational activation of ErbB2 reveals a new protein kinase autoinhibition mechanism. J Biol Chem. 2008; 283:1588–1596. [PubMed: 18039657]
61. Moore MJ, et al. Structural basis for peptide binding in protein kinase A. Role of glutamic acid 203 and tyrosine 204 in the peptide-positioning loop. J Biol Chem. 2003; 278:10613–10618. [PubMed: 12499371]
62. Yang J, et al. Allosteric network of cAMP-dependent protein kinase revealed by mutation of Tyr204 in the P+1 loop. J Mol Biol. 2005; 346:191–201. [PubMed: 15663937]
63. Masterson LR, et al. Allosteric cooperativity in protein kinase A. Proc Natl Acad Sci U S A. 2008; 105:506–511. [PubMed: 18178622]
64. Yang J, et al. A conserved glu-arg salt bridge connects coevolved motifs that define the eukaryotic protein kinase fold. J Mol Biol. 2012; 415:666–679. [PubMed: 22138346]
65. Dar AC, et al. Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR. Cell. 2005; 122:887–900. [PubMed: 16179258]
66. Kim C, et al. PKA-I holoenzyme structure reveals a mechanism for cAMP-dependent activation. Cell. 2007; 130:1032–1043. [PubMed: 17889648]
67. Song HW, et al. Phosphoprotein-protein interactions revealed by the crystal structure of kinase-associated phosphatase in complex with PhosphoCDK2. Molecular Cell. 2001; 7:615–626. [PubMed: 11463386]
68. Hantschel O, et al. A myristoyl/phosphotyrosine switch regulates c-Abl. Cell. 2003; 112:845–857. [PubMed: 12654250]
69. Deminoff SJ, et al. Distal recognition sites in substrates are required for efficient phosphorylation by the cAMP-dependent protein kinase. Genetics. 2009; 182:529–539. [PubMed: 19364808]
70. Masterson LR, et al. Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. Proc Natl Acad Sci U S A. 2011; 108:6969–6974. [PubMed: 21471451]
71. Masterson LR, et al. Allostery and binding cooperativity of the catalytic subunit of protein kinase A by NMR spectroscopy and molecular dynamics simulations. Adv Protein Chem Struct Biol. 2012; 87:363–389. [PubMed: 22607761]
72. Srivastava AK, et al. Synchronous Opening and Closing Motions Are Essential for cAMP-Dependent Protein Kinase A Signaling. Structure. 2014; 22:1735–1743. [PubMed: 25458836]
73. Prowse CN, Lew J. Mechanism of activation of ERK2 by dual phosphorylation. J Biol Chem. 2001; 276:99–103. [PubMed: 11016942]
74. Yang J, et al. Crystal structure of a cAMP-dependent protein kinase mutant at 1.26 angstrom: New insights into the catalytic mechanism. Journal of Molecular Biology. 2004; 336:473–487. [PubMed: 14757059]
75. Bastidas AC, et al. Influence of N-myristylation and ligand binding on the flexibility of the catalytic subunit of protein kinase A. Biochemistry. 2013; 52:6368–6379. [PubMed: 24003983]
76. Zhang J, et al. Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors. Nature. 2010; 463:501–506. [PubMed: 20072125]

77. Piserchio A, et al. Assignment of backbone resonances in a eukaryotic protein kinase - ERK2 as a representative example. Methods Mol Biol. 2012; 831:359–368. [PubMed: 22167683]
78. Palmer AG 3rd. NMR characterization of the dynamics of biomacromolecules. Chem Rev. 2004; 104:3623–3640. [PubMed: 15303831]
79. Kim J, et al. Dysfunctional conformational dynamics of protein kinase A induced by a lethal mutant of phospholamban hinder phosphorylation. Proc Natl Acad Sci U S A. 2015; 112:3716–3721. [PubMed: 25775607]
80. Olah GA, et al. Solution structure of the cAMP-dependent protein kinase catalytic subunit and its contraction upon binding the protein kinase inhibitor peptide. Biochemistry. 1993; 32:3649–3657. [PubMed: 8385485]
81. Herberg FW, et al. Dissection of the nucleotide and metal-phosphate binding sites in cAMP-dependent protein kinase. Biochemistry. 1999; 38:6352–6360. [PubMed: 10320366]
82. Li F, et al. Evidence for an internal entropy contribution to phosphoryl transfer: a study of domain closure, backbone flexibility, and the catalytic cycle of cAMP-dependent protein kinase. J Mol Biol. 2002; 315:459–469. [PubMed: 11786025]
83. Wand AJ, et al. A surprising role for conformational entropy in protein function. Top Curr Chem. 2013; 337:69–94. [PubMed: 23478875]
84. Xiao Y, et al. Phosphorylation releases constraints to domain motion in ERK2. Proc Natl Acad Sci U S A. 2014; 111:2506–2511. [PubMed: 24550275]
85. Tirion MM. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. Phys Rev Lett. 1996; 77:1905–1908. [PubMed: 10063201]
86. Mahajan S, Sanejouand YH. On the relationship between low-frequency normal modes and the large-scale conformational changes of proteins. Arch Biochem Biophys. 2015; 567:59–65. [PubMed: 25562404]
87. Shudler M, Niv MY. BlockMaster: partitioning protein kinase structures using normal-mode analysis. J Phys Chem A. 2009; 113:7528–7534. [PubMed: 19485335]
88. Bhattacharyya M, Vishveshwara S. Probing the allosteric mechanism in pyrrolysyl-tRNA synthetase using energy-weighted network formalism. Biochemistry. 2011; 50:6225–6236. [PubMed: 21650159]
89. Sethi A, et al. Dynamical networks in tRNA:protein complexes. Proc Natl Acad Sci U S A. 2009; 106:6620–6625. [PubMed: 19351898]
90. Kappel K, et al. The binding mechanism, multiple binding modes, and allosteric regulation of Staphylococcus aureus Sortase A probed by molecular dynamics simulations. Protein Sci. 2012; 21:1858–1871. [PubMed: 23023444]
91. Vijayabaskar MS, Vishveshwara S. Interaction energy based protein structure networks. Biophys J. 2010; 99:3704–3715. [PubMed: 21112295]
92. James KA, Verkhivker GM. Structure-based network analysis of activation mechanisms in the ErbB family of receptor tyrosine kinases: the regulatory spine residues are global mediators of structural stability and allosteric interactions. PLoS One. 2014; 9:e113488. [PubMed: 25427151]
93. Palla G, et al. Uncovering the overlapping community structure of complex networks in nature and society. Nature. 2005; 435:814–818. [PubMed: 15944704]
94. McClendon CL, et al. Dynamic architecture of a protein kinase. Proc Natl Acad Sci U S A. 2014; 111:E4623–E4631. [PubMed: 25319261]
95. Girvan M, Newman ME. Community structure in social and biological networks. Proc Natl Acad Sci U S A. 2002; 99:7821–7826. [PubMed: 12060727]
96. McClendon CL, et al. Quantifying Correlations Between Allosteric Sites in Thermodynamic Ensembles. J Chem Theory Comput. 2009; 5:2486–2502. [PubMed: 20161451]
97. Kannan, N., et al. Analogous regulatory sites within the alpha C-beta 4 loop regions of ZAP-70 tyrosine kinase and AGC kinases. Elsevier Science Bv; 2008. p. 27-32.
98. Hanks SK, et al. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science. 1988; 241:42–52. [PubMed: 3291115]
99. Knighton DR, et al. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science. 1991; 253:407–414. [PubMed: 1862342]

100. Changeux JP. Allostery and the Monod-Wyman-Changeux model after 50 years. Annu Rev Biophys. 2012; 41:103–133. [PubMed: 22224598]

101. Bhattacharyya RP, et al. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu Rev Biochem. 2006; 75:655–680. [PubMed: 16756506]

102. Lockless SW, Ranganathan R. Evolutionarily conserved pathways of energetic connectivity in protein families. Science. 1999; 286:295–299. [PubMed: 10514373]

103. Suel GM, et al. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. Nat Struct Biol. 2003; 10:59–69. [PubMed: 12483203]

104. Poulikakos PI, et al. RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF. Nature. 2010; 464:427–430. [PubMed: 20179705]

105. Berman HM, et al. The Protein Data Bank. Nucleic Acids Res. 2000; 28:235–242. [PubMed: 10592235]

Box 1

"Domino" and "Violin" models of signal transduction

When a small molecule binds to the allosteric site of a protein, information is transferred through the protein molecule to its active site. Two different ways of such transmission can be defined. The first mechanism here defined as the "Domino model" is a sequential set of events propagating linearly from the allosteric site to the active site. Binding of the effector triggers local structural changes that sequentially propagate via a single pathway to the active site. It was suggested that this mechanism is applicable for the PDZ domain family [102], G-protein coupled receptors, the chymotrypsin class of serine proteases and hemoglobin [103]. The second mechanism, defined here conceptually as a "Violin model", is based on vibration pattern changes inside the protein. In a violin its pitch can be changed by a slight movement of the violin player's finger on the fingerboard. Information about the finger movement is, thus, transferred throughout the whole body of the violin with no specific pathway for the signal transduction. By analogy, protein allosteric site is a fingerboard of the protein and a small signaling molecule is the player's finger. If a protein is in a particular vibration mode, it is possible to suggest that binding a small effector molecule to a specific site can change this mode. The signal, thus, will be spread throughout the whole protein including its active site. The "Domino model" is a reliable way to transfer information in a macro world, but on a molecular level, with significant thermal motions of the protein, this mechanism will be prone to random triggering of the domino chain reaction, creating noise in the signaling system. Thermal motions in case of the "Violin model" do not hinder the transduction. In fact, the permanent motion of the molecule is a prerequisite for this mechanism.

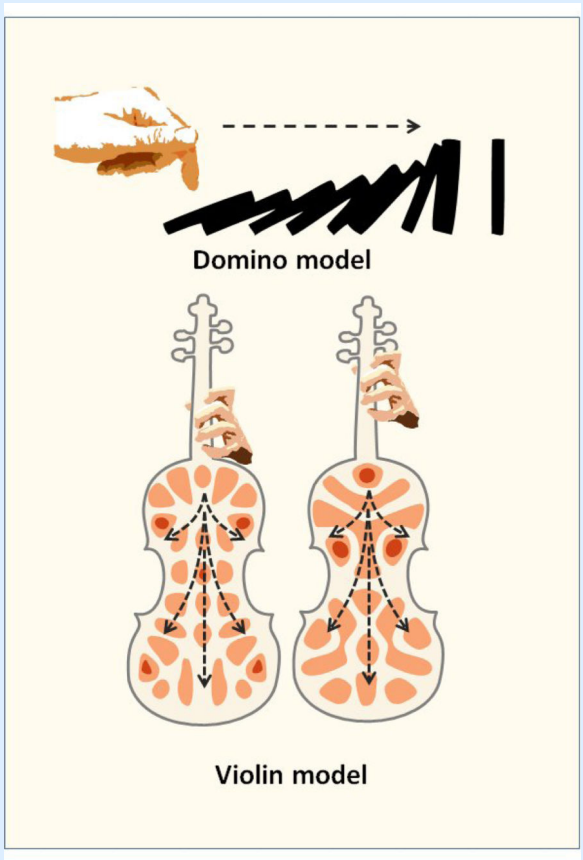

Figure I. Two different mechanisms of signal transduction.

Box 2

Critical role of αC-β4 Loop in different kinases

Although the geometry of the αC-β4 Loop in active EPKs is highly conserved and does reveal significant changes in activation process, it was implicated in many regulatory events. In BRAF the conserved Tyr566 (Tyr156 in PKA) positions main chain of Arg509 (**Figure 1A**) that is important for activation of this kinase through dimerization[104]. The dimer interface involves the αC-β4 Loops of both oligomers. The analogous arginine from the binding partner makes a set of strong hydrogen bonds in the αC-β4 region and serves as a capping residue for the αC-helix (**Figure 1B**). Similar geometry and interactions with the conserved tyrosine can be observed in the kinase domain of EGFR (**Figure 1C**). A close homolog of EGFR, ErbB2, is has two unusual substitutions: Gly776 and Gly778, that are a serine and an aspartate in EGFR (**Figure 1C**). ErbB2, unlike EGFR, reveals no significant in vitro kinase activity despite the correct assembly of the Regulatory spine. The ErbB2 structure shows that the αC-β4 Loop of ErbB2 has a significantly different geometry. The hydrogen bond to the αE-helix tyrosine is broken. Mutation of these glycines to serine and aspartate restores catalytic activity, suggesting that excess flexibility of the αC-β4 Loop is detrimental to the kinase activity [60].

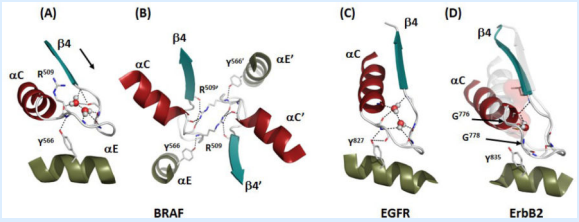

Figure I. αC-β4 Loop interactions in BRAF, EGFR and ErbB2. αC-helices are shown in dark red, αE-helices in olive and β4-starnds in teal. Water molecules are shown as red and white spheres. **(A,B)** BRAF (PDBID: 4E26). **(C)** EGFR (PDBID: 4WRG). **(D)** ErbB2 (PDBID: 3PP0). EGFR structure is overlaid for comparison and shown as transparent cartoon. R-spine for EGFR is shown as transparent red surface.

Outstanding questions

- Do different active protein kinases have similar community maps?
- How are community maps changed or regulated by protein:protein interactions such as dimer activation and binding of protein substrates?
- How are communities defined in inactive kinases?
- Can inactive protein kinases be classified in terms of dynamics instead of traditional “DFG-in/DFG-out” categories?
- How do disease mutations alter the communities of the kinase core either negatively or positively

Trends Box

- Allostery in protein kinases is based on changes in dynamics of large ensembles of residues that emerge stochastically inside the protein.
- Active kinases are in a particular dynamic mode that is characterized by an ATP-dependent μs-ms timescale that radiates across the molecule.
- The protein kinase core consists of dynamic modules (communities) that move as semirigid bodies and correlate with catalytic and regulatory functions. Non-conserved elements such as tails and linkers are integrated with the core communities and are essential for the dynamic regulation of the core.
- Assembly of the active kinase, as defined by the integrated communities, is mediated by the assembly of the regulatory spine and often facilitated by Activation Loop phosphorylation.
- Dynamics mediate the classic connection between a protein structure and its function even when major conformational changes are not apparent.

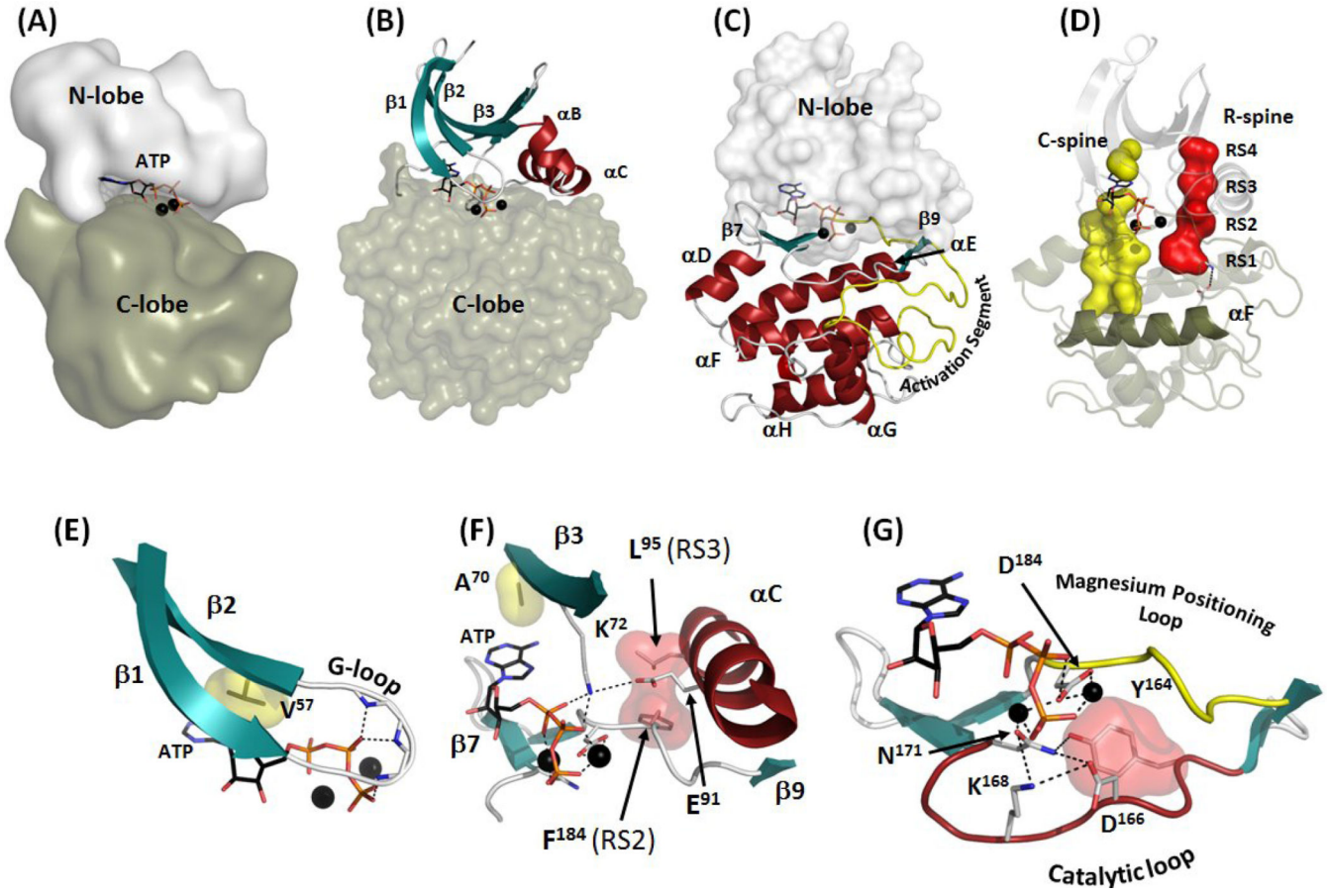

Figure 1.

Internal architecture of the protein kinase core. (**A**) Bilobal structure of protein kinase core shows ATP imbedded in the deep cleft between the smaller N-lobe (white surface) and the larger C-lobe (olive surface). (**B**) The N-lobe has five β-strands (teal) and a large α-C helix (red). (**C**) The C-lobe is predominantly helical with a large Activation Segment between β8 and αF (yellow). (**D**) Two “hydrophobic spines” are assembled in active kinases: The R-spine (red surface) that contains four residues (RS1 through RS4) is bound to the αF-helix via conserved hydrogen bond and is correctly assembled only in the active kinase. The C-spine (yellow surface) is completed by adenine ring of ATP. The spines connect the two lobes and span across the molecule from the large αF-helix in the C-lobe to the rigid β-sheet in the N-lobe. (**E**) Glycine-rich Loop (G-loop) lies between β1 and β2 and coordinates the phosphate groups of ATP. β2-strand also contains the conserved Val57 that is a part of the C-spine. (**F**) β3-strand contains two universally conserved residues, Ala70 and Lys72. The former is a part of the C-spine and the latter coordinates the α- and β-phosphates of ATP and binds to the invariant Glu91 from the αC-helix, stabilizing the R-spine (red surface) and ordering the αC-helix for catalysis. (**G**) Residues that are directly involved in the phosphotransfer are in the Catalytic loop between β6 and β9 (red) with Asp184 coming from the Magnesium Positioning Loop (yellow).

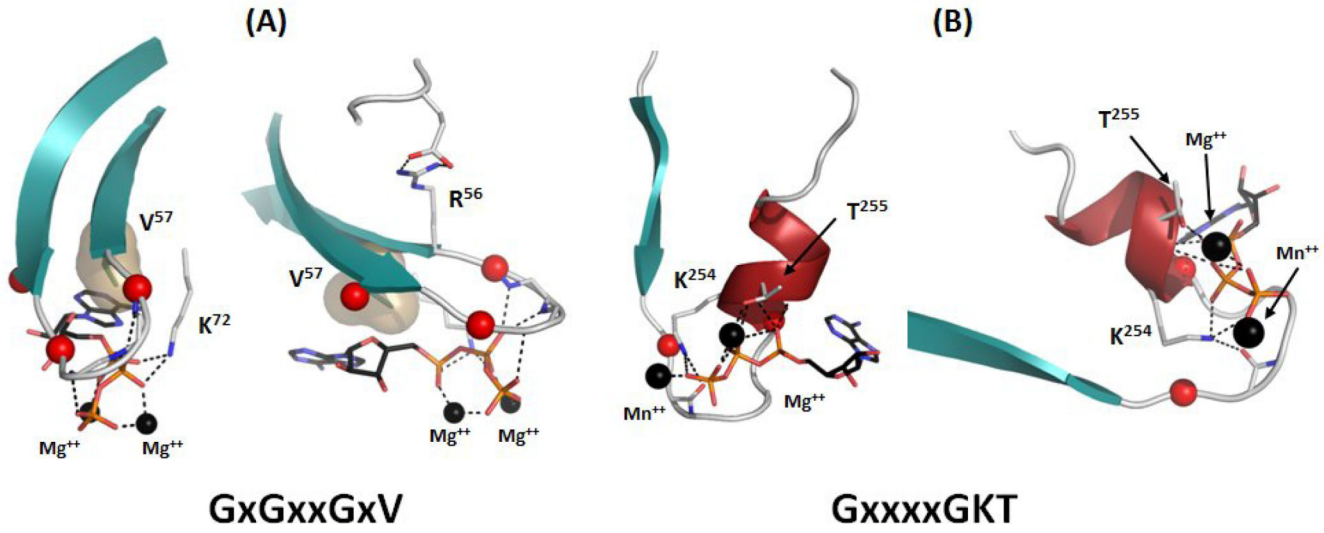

Figure 2.

G-loops and P-loops. The G-loop of PKA (Protein Data Bank[105] identifier (PDBID): 1ATP) is compared to the P-loop of Phosphoenolpyruvate Carboxykinase (PCK) (PDBID: 1AQ2). (**A**) The G-loop connects two β-strands and contains three glycines (red spheres). The residue that follows the last glycine points away from the ATP phosphates. Right after this residue there is a highly conserved valine that is a part of the C-spine (tan surface). (**B**) P-loop connects a β-strand and an α-helix and contains two conserved glycines (red spheres) that are followed by lysine and serine/threonine. Unlike Arg56 and Val57 in PKA Thr255 and Lys254 in PCK contribute directly to metal and phosphate binding.

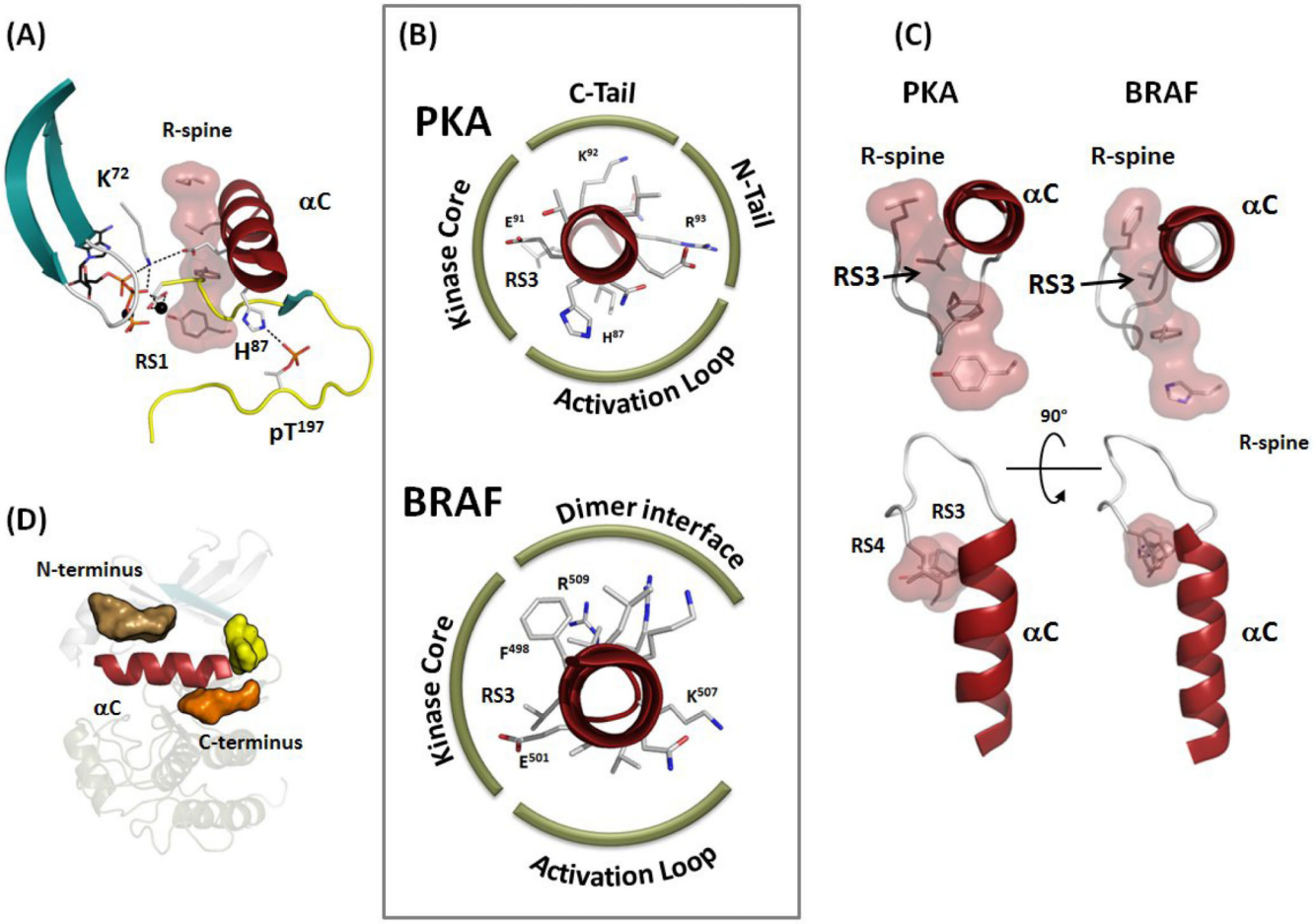

Figure 3.

Multiple ways to stabilize the αC-helix. (**A**) N-terminal part of the αC-helix is often bound to the phosphorylated residue in the Activation Loop, while the middle part of the helix is stabilized by the conserved Lys72-Glu91 salt bridge. (**B**) αC-helices of PKA and BRAF shown as a wheel highlighting their interactions. While interactions of the αC-helix with the kinase core is usually conserved, other surfaces of the helix have different binding partners. (**C**) C terminus of the αC-helix in all kinases is anchored to the kinase core by the αC-β4 Loop and via integration of its RS3 residue with the R-spine. (**D**) Conserved pockets flank the surface of the αC-helix as detected by Thompson et al. [46]. One pocket is located at the N terminus (tan surface) and two pockets at the C terminus (yellow and orange surfaces).

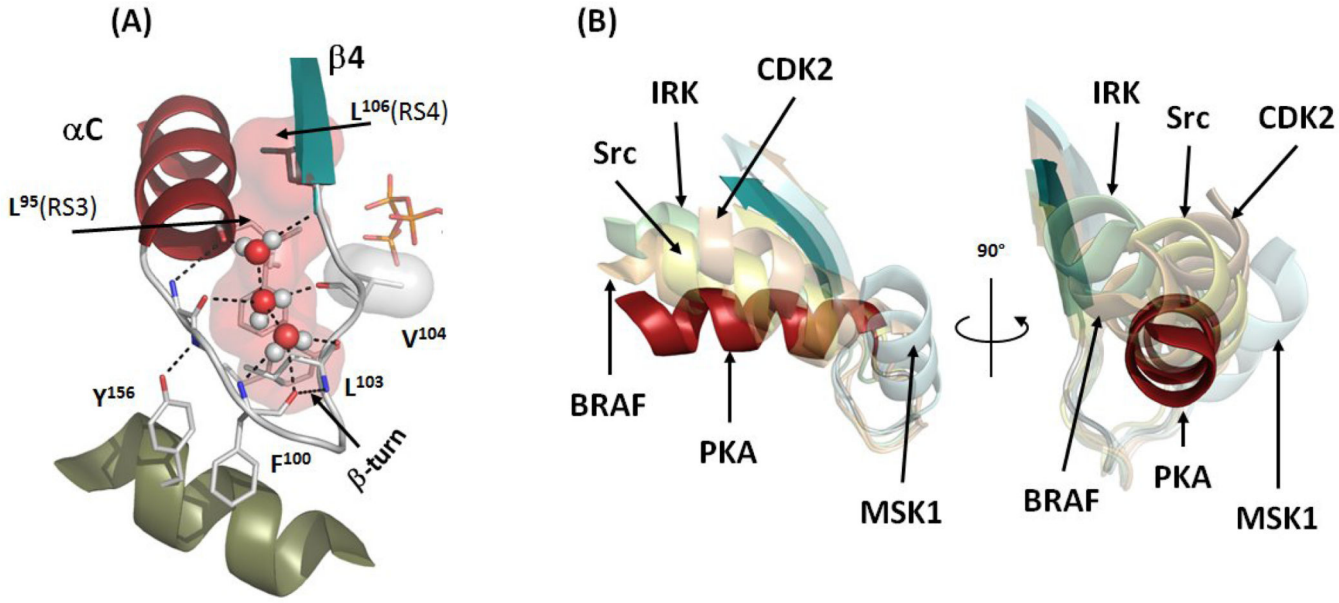

Figure 4.

αC-β4 Loop is a pivot point for the αC-helix movement. (**A**) The αC-β4 Loop (white) connects aC-helix (red) and β4 (teal). It is anchored to the αE-helix (olive) via interaction with conserved Tyr156. The geometry of the loop is highly conserved with the β-turn between Phe100 and Leu103 and three water molecules (red and white spheres). The assembled R-spine is shown as a red surface. Val104 is shown as a white surface. (**B**) The αC-β4 Loops of five inactive kinases (transparent cartoons) in comparison with the αC-β4 Loop of active PKA (red and teal). The structures were aligned by their C-lobes. The following structures were used: PKA (PDBID:1ATP), BRAF (PDBID:1UWH), Src (PDBID:2SRC), IRK (PDBID:1IRK), CDK2 (PDBID:1B39), MSK1 (PDBID:1VZO).

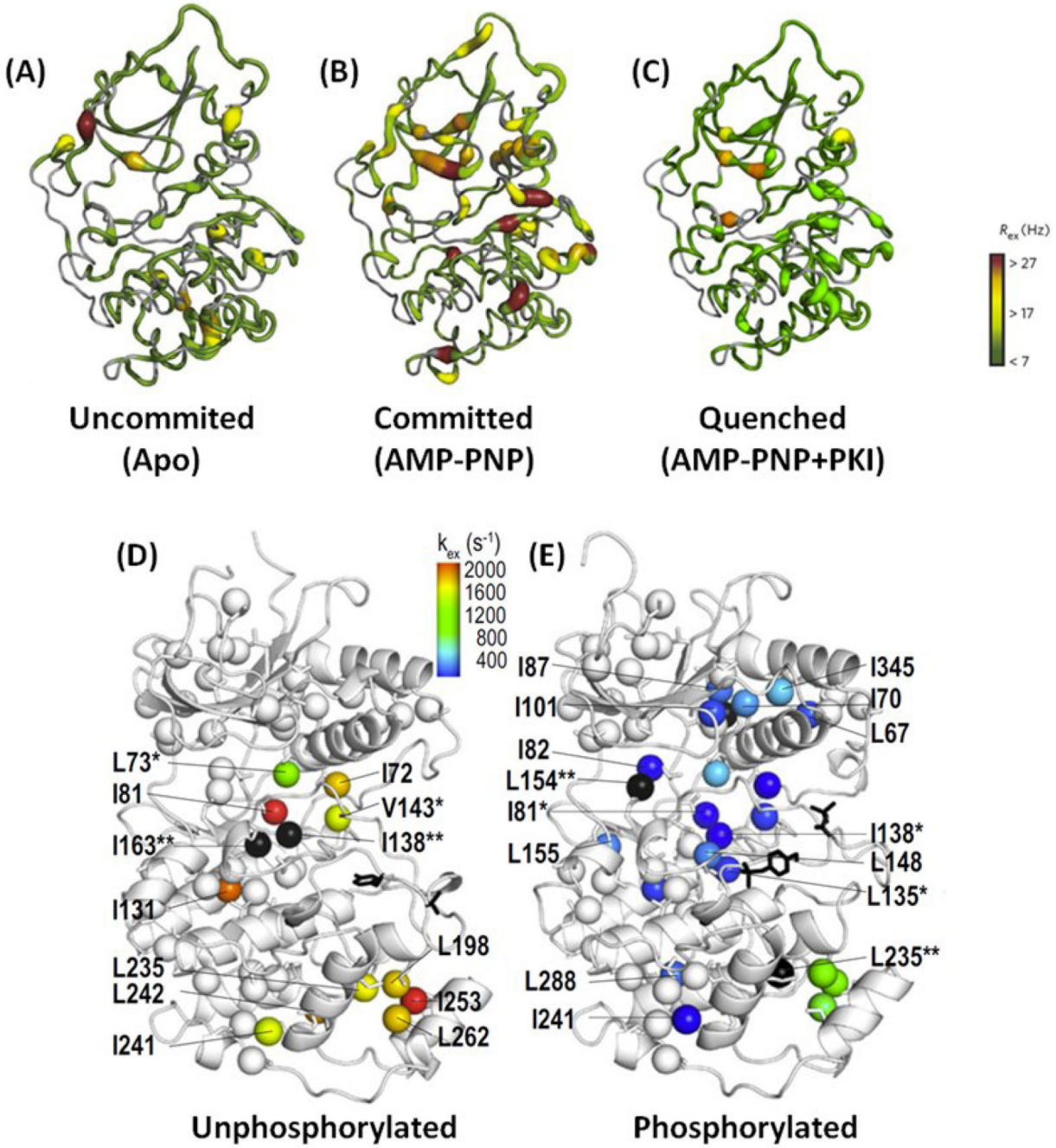

Figure 5.

Changes in slow dynamics in PKA and ERK2. (**A**) Nuclear spin relaxation rate (R_{ex}), that reflects amino acid dynamics in μs-ms timescale, is mapped onto the PKA structure. Relatively low signal is detected in PKA with no nucleotide bound. (**B**) Binding of the ATP-analog AMP-PNP causes a significant increase of R_{ex} throughout the whole molecule. (**C**) Binding of peptide inhibitor (PKI) leads to significant decrease of the slow dynamics. (Adapted from [26]) (**D**) Relaxation dispersion analysis of [^{13}C]-labeled methyl groups provides information on dynamics in the μs-ms range in ERK2. Not phosphorylated form of

ERK2 reveals diverse dynamic properties of hydrophobic sidechains in the range of 1–2kHz. (**E**) Double phosphorylation of the Activation Loop causes a significant decrease in the dynamics rate and leads to a more uniform dynamic profile of the kinase core. (Adapted from [84])

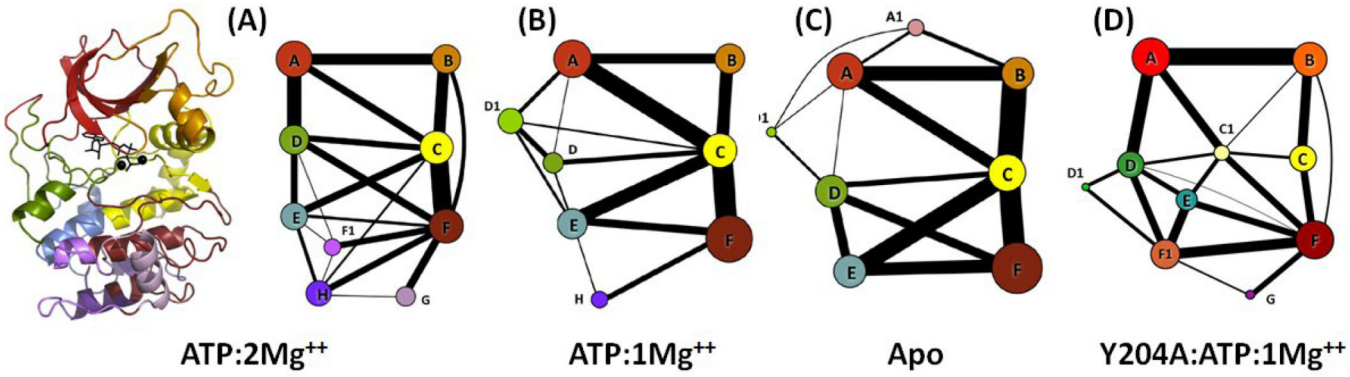

Figure 6.

Community map analysis. (**A**) Nine communities detected in PKA with ATP and two magnesium ions bound mapped onto PKA structure (left) and presented as a graph (right). Vertex size is proportional to the number of residues in the community. Edge widths reflect the degree of communication between communities. (**B**) Community map for PKA bound to ATP and one magnesium ion. (**C**) Community map for the Apo form of PKA (**D**) Community map of Y204A mutant of PKA bound to ATP and one magnesium ion. Adapted from [94] and [72].

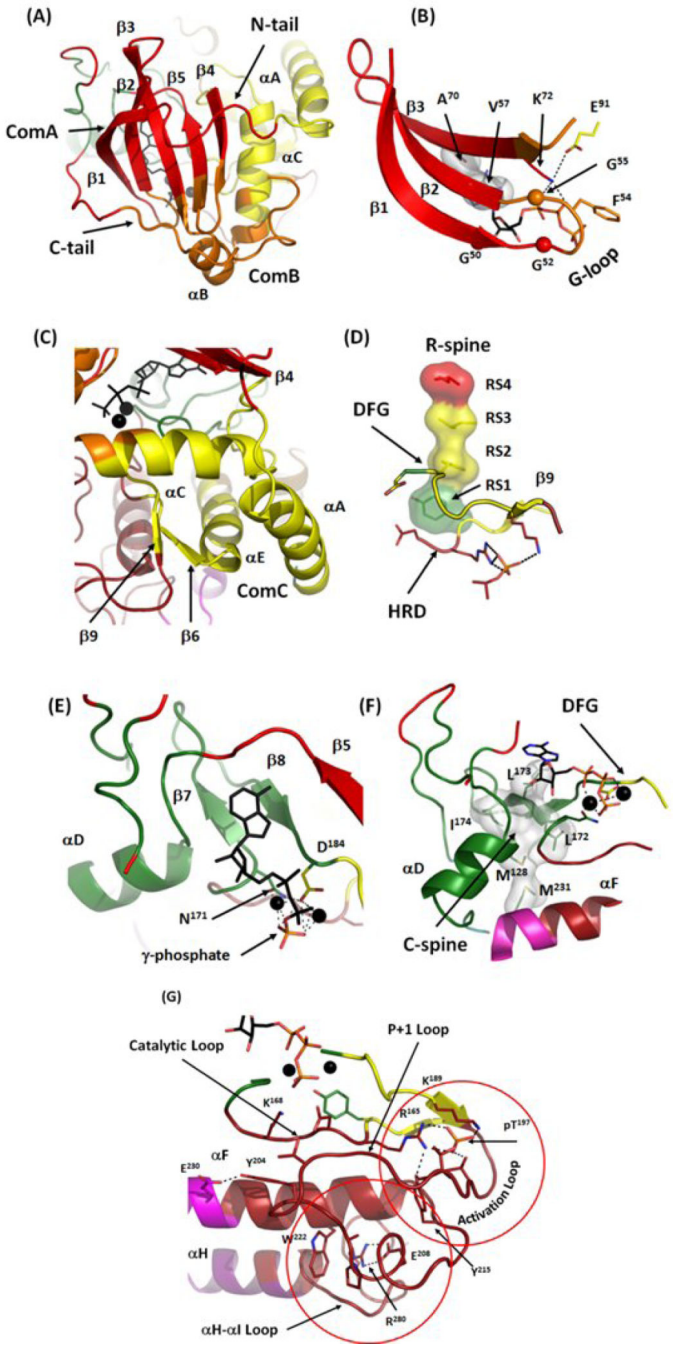

Figure 7.

Detailed view of major communities in active PKA. (**A**) Community A (red) includes most of the rigid β-sheet from the N-lobe and adjacent parts of N- and C-tails. Community B (orange) includes the end of the C-tail, the αB-helix, a part of the αC-helix and two loops the G-loop and β4-β5. (**B**) Close-up of the community assignments around the G-loop. The three conserved glycines are shown as spheres. Most of the G-loop is associated with Community B with Phe54 facing the αB-helix. ATP binding is mostly controlled by the residues from Community A: V57 and Ala70 from the C-spine and Lys72 from the β3, but

the β- and γ-phosphates of ATP, are positioned by Community B. (**C**) Community C (yellow) includes most of the αC-helix, a large portion of the α-E helix, the β6-β9 sheet and the N-tail of PKA with the αA-helix. (**D**) Community C contains elements that are the most critical for active conformation: the DFG-motif, RS3 from the αC-helix and the β6-β9 sheet. (**E**) Community D (green). Adenine ring and ribose of ATP are supported by the β7-β8 sheet including three C-spine residues on β7 (Leu172, Leu173 and Ile174). Community D also includes the most of important residues for catalysis and magnesium coordination, the main chain of Asp184 from the DFG-motif and Asn171 from the Catalytic Loop. (**F**) Most of the C-spine (white surface) is a part of the Community D. (**G**) Community F. This community includes the EPK-specific Activation Segment, the Catalytic Loop, most of the αF-helix and the αH-αI Loop. It is built around two conserved structural elements (red circles): the phosphorylated residue in the Activation Loop (Thr197) and the buried salt bridge between Glu208 from the APE-motif and Arg280 from the αH-αI Loop.

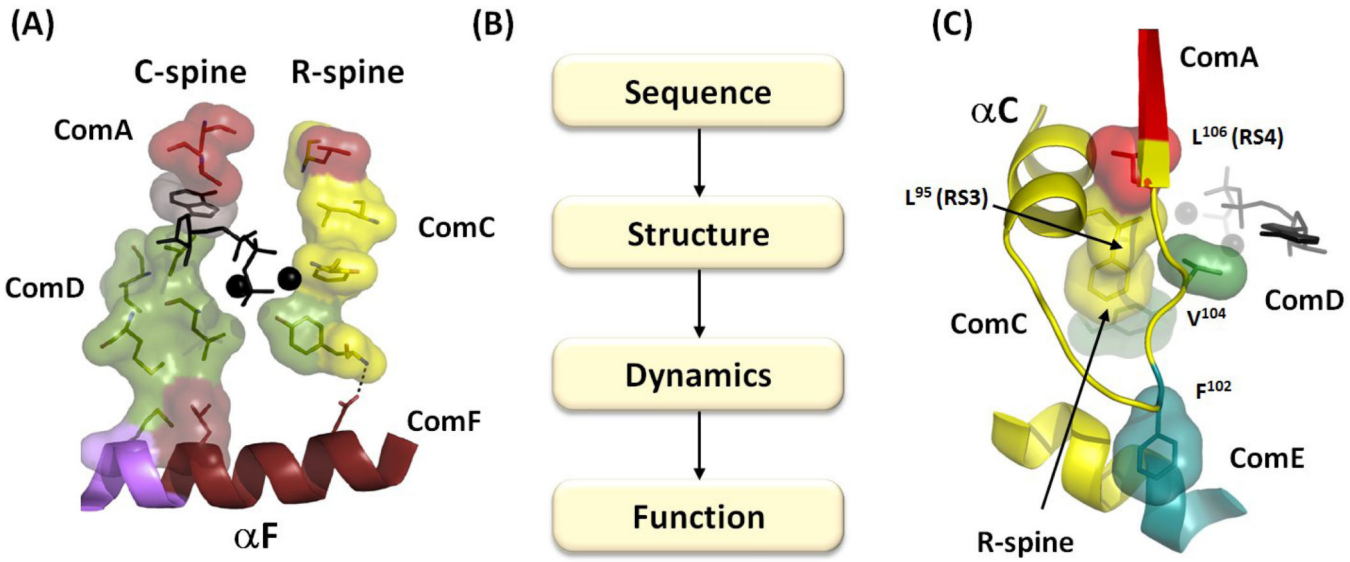

Figure 8. Key figure

Protein dynamics define communication lines inside the kinase core. (**A**) Two hydrophobic spines span the kinase core and connect major communities that are responsible for catalysis and regulation: A, C, D and F. (**B**) The amended central dogma of biology where the connection between protein structure and its function is mediated by protein dynamics. (**C**) The conserved αC-β4 Loop appears to be an important "communication hub" that connects Community A from the N-lobe to three C-lobe communities: Community D, which plays an important role in catalysis, regulatory Community C and Community E. It includes the two R-spine residues in the N-lobe (RS3 and RS4) and anchors them to the two R-spine residues from the C-lobe when the active conformation is assembled.