Insights into evolutionary interaction patterns of the 'Phosphorylation Activation Segment' in kinase

Adil Ahiri1, Hocine Garmes2, Crtomir Podlipnik3, Aziz Aboulmouhajir1,4

1Modeling and Molecular Spectroscopy Team, Faculty of Sciences, University Chouaib Doukkali, El-Jadida, Morocco; 2Analytical Chemistry and Environmental Sciences Team, Department of chemistry, Faculty of Science, University Chouaib Doukkali, El Jadida, Morocco; 3Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia; 4Extraction, Spectroscopy and Valorization Team, Organic synthesis, Extraction and Valorization Laboratory, Faculty of Sciences of Ain Chock, Hassan II University, Casablanca, Morocco. *Corresponding authors email: aboulmouhajir@gmail.com & ahiri.adil@gmail.com

Received June 15, 2019; Revised October 7, 2019; Accepted October 9, 2019; Published October 13, 2019

DOI: 10.6026/97320630015666

Abstract:
We are interested in studying the phosphorylation of the kinase activation loop, distinguishing the passage from the unphosphorylated to the phosphorylated form without allostery. We performed an interaction study to trace the change of interactions between the activation segment and the kinase catalytic core, before and after phosphorylation. Results show that the structural changes are mainly due to the attraction between the phosphate group and guanidine groups of the catalytic loop, the β9, and the αC helix. This attraction causes propagation of structural variation of the activation segment, principally towards the N-terminal. The structural variations are not made on all the amino acids of the activation segment; they are conditioned by the existence of two beta sheets stabilizing the loop during phosphorylation. The first, β6-β9 sheet is usually present in most of the kinases; the second, β10-β11 is formed due to the interaction between the main chain amino acids of the activation loop and the αEF/αF loop.

Keywords: Kinase, activation segment, phosphorylation, structural variation, interaction variation.

Background:
The reversible interplay between protein kinases and Phosphatases has an important role in cellular signalization [1]. Protein kinases are enzymes which mediate the transfer of γ-phosphate group (PO42−) from ATP to target protein substrates [1]. On the contrary, phosphatases remove phosphate groups from the phosphorylated protein substrate [1]. The protein kinases have short consensus sequences in common, which define the catalytic subunit of about 270 amino acids, distributed in two lobes: the N terminal lobe (about 80 residues) consisting essentially of a β-sheet with five strands and a Cα helix [2]. The loop connecting the first two strands β is P-loop. The C-terminal lobe (about 190 residues) consists mostly of α-helices and two essential close loops, which are catalytic loop and activation loop [2]. The latter belongs to the activation segment which runs from the conserved DFG (Asp-Phe-Gly) in the magnesium binding loop to the conserved APE (Ala-Pro-Glu) and includes β9, the activation loop and the P+1 loop, moving from N-terminal to C-terminal anchor points [3]. The conformation of the whole activation segment controls kinase activity [3]. The two lobes are connected by a so-called hinge region ensuring their mutual flexibility [2]. The phosphorylation of the substrate occurs in the catalytic site or active site, which is at the intersection of N and C lobes (Figure 1). This site contains two juxtaposed pockets to receive the ATP and the substrate, respectively [2]. The regulation by phosphorylation of the kinase is essential for its catalytic competence [3]. It can take place either directly by modulating the ATP binding, or the substrate binding
on their catalytic site indirectly, by the displacement, to blocking key elements of the catalytic, or regulation domain [3]. The kinase can also be catalytically inactive due to the displacement of key elements participating in the catalytic domain, which is then allosteric regulation [3]. The activation of some kinases requires phosphorylation of their activation segment, particularly their activation loop which is a real phosphorylation regulatory site [3]. The activation loop is involved in the kinase regulations which gives an essential role in the adoption of an active conformation. SLK phosphorylation at T183 and S189 levels is known [4]. Phosphorylation plays an important role in the activation and signalling processes of SLK [4]. Besides, the regulatory effect extends to the catalytic environment by the activation segment. Researchers have recorded a deregulation of MAPK at low pH due to a structural rearrangement of the activation segment [5]. The latter may also have an auto-inhibitory action because of substrate blocking and stabilisation of an inactive conformation of the αC helix, such as the case of NDRI in its non-phosphorylated state [6]. What is more, one of the most important mechanisms related to the activation segment is auto phosphorylation, where kinases dimerize and cause phosphorylation of the activation segment [7, 8, 9]. So, the exchange of activation segment has proven to be a very important action in the auto phosphorylation mechanism [11], which can occur in one or both directions [12]. It is remarkable that autophosphorylation does not recognize the consensus sequences of substrates, which gives us a unique phosphorylation mechanism [10]. Concerning the physical aspect, the activation segment starts from the conserved motif DFG, which is the part of the activation loop, to the conserved motif APE (Figure 1) [3]. The N-terminal and C-terminal anchor points of the activation segment take place despite the folding of the activation loop under the effect of the phosphorylation of one of the kinase residues [3]. In fact, in many kinases, the interaction of the phosphorylated residue with the activation loop gives two states. One concerns the aspartate of the DFG motif pointing in the ATP binding site and coordinates two Mg²⁺ ions (active state) DFGin, and the other concerns DFG pointing out of the ATP binding site (inactive state, DFGout) [13]. When the N-terminal or the C-terminal anchor points are disturbed, the kinase is in an inactive state. In general, it is commonly accepted that in the cell environment, kinases pass between the catalytically active conformation and inactive conformation [13]. There are also other secondary phosphorylation sites, both upstream and downstream of the primary site, improving the regulation of kinase activity [3]. Protein kinases are subdivided into subfamilies according to their different catalytic domains specificity and to target amino acids, knowing that tyrosine phosphorylation has attracted more interest in biomedical research thanks to its relation to human disease via the dysregulation of receptor tyrosine kinases (RTKs) [14]. Indeed, currently, 37 kinase inhibitors have received FDA approval for the treatment of cancer, and about 150 kinase-targeted drugs are in various clinical phase trials [14]. The human protein kinases numbering 518, constitute one of the largest family of the human genome, representing 1.7% of the genome [15]. These protein kinases are subdivided into two categories: the first, with 478 families, has typical eukaryotic catalytic domains (ePKs), and the second, contains 40 families with atypical catalytic domains [16]. The study of sequence identities of the catalytic domains of ePKs revealed 491 catalytic domains divided into nine groups [17]. Frequently, the non-catalytic domains are misnamed as regulatory domains, although the regulation and phosphorylation can also take place in the catalytic domain in this case [3]. Several structural studies have reported the interaction between the phosphorylated amino acid and the catalytic body without integrating the whole activation segment [3]. We report here the correlation of the “phosphorylation activation segment” P.A.S and more specifically activation loop, with principal residues in kinase activity. We further explain the structural changes due to the P.A.S - kinase catalytic core interaction, benefiting from increasingly abundant databases of protein structures, and avoiding expensive quantum or dynamic calculations. Hence, a differential structural analysis between phosphorylated and non-phosphorylated P-kinases (kinase catalytic core) is completed.

Figure 1: The P-kinase domain of GSK3B showing the activation segment and its various constituents.

Materials and methods:
Dataset preparation:
The P-kinases sequences were extracted from the UNIPROT database [18], and their 3D structures were identified in the PDB RCSB database [19] by BLASTP. During alignment with BLAST, we have chosen an E-value=0 to avoid all the alignments randomly and to be able to extract structures principally with identical sequences from the UNIPROT database [20]. The PhosphoSite database [21] allows us to locate the phosphorylated amino acids in kinases sequences. We have retained only structures with resolved Activation segments, and their visualization was made in the Chimera program [22]. For a successful comparison between phosphorylated and non-phosphorylated kinases, it was necessary to build a homogeneous database, thereby removing different factors that may interfere with phosphorylation. Therefore, a selectivity protocol integrating three filters was set up. The first filter consists of comparing P-kinases of the same stoichiometry [23]. While the second filter annihilates the allosteric effect in the active sites, by eliminating all structures having ligands positioned in their allosteric sites [24]. We selected all the ligands possessing the best value of interaction fingerprint compared with ATP for further studies [25]. The third and last filter, called the RMSD (Root-Mean-Square Deviation) Cα clustering effect, which measures the main chain carbons fluctuations, has the role of avoiding, among the structures proposed in the PDB for the same kinase, all of the very fluctuating structures giving rise to a Cα RMSD exceeding 2 Å during their comparison [26].

Dataset treatment:
After the selection of the phosphorylated and not phosphorylated structures by the different filters, we focused only on the catalytic domain, using the annotation provided by PROSITE [27]. Then, the missing sequences of this domain, which are not solved by XRAY, are modelled by MODELLER on the basis on PDB of the same class previously collected during preparation [28] and validated (checked) by Meta server SAVES. The next step consists of rectifying mutations using the most crucial score of Dynamics Rotamer Library [29].

Dataset analysis:
We have studied the conservation of secondary structures in the activation segment for selected P-kinases, before discussing the difference between compared 3D structures of the unphosphorylated structure considered as a reference, and phosphorylated structure for the extraction of the phosphorylation effect. For this, alignment of sequences, and a superposition of their secondary structures have been carried out.

We have performed an analysis of the cartesian deviations of the backbone chain by RMSDbb, side chain, by RMSDsc, and also the deviations of the dihedral angles ϕ (φ), ψ (ψ) of each amino acid of our kinase to detect structural disparities between the phosphorylated and non-phosphorylated activation loop structures [30]. We have retained only variations which exceed 2 Å for RMSDbb and RMSDsc [31] and 20° for the dihedral angles [32]. Finally, to identify the impact of phosphorylation of the activation loop on the modes of its interaction with the remaining P-kinase motifs, we have attempted to enumerate the amino acids of activation loop which are essential for such interactions, using the FICI script [33].

| GROUP | KINASE (FAMILY) | Stochiometry | ID | Species | Ligands | MUTATION |
|-------|----------------|--------------|----|---------|---------|----------|
| CMGC  | CDK2 (CDK)     | AB           | 450Q_A | Homo sapiens | ATP, MG, TPO<sub>4</sub> |
|       |                | AB           | 1FN_A | Homo sapiens | ATP     |
|       |                | A5           | 2OK_A | Homo sapiens | IBM     |
|       |                | A5           | 2FW3_A | Homo sapiens | BMI, PTR<sub>4</sub> |
|       |                | A5           | TP8_A | Mus musculus |         |
|       |                | A5           | SPY_A | Mus musculus | PTR<sub>4</sub>, TPO<sub>4</sub> |
|       |                | A5           | 28Y_A | Homo sapiens | ATP, GOL, SO4 |
|       |                | A5           | 3RC2_A | Homo sapiens | 3RC, SEP<sub>4</sub> |
|       |                | A5           | 1VT2_A | Homo sapiens | 4DG, GOL, PTR<sub>4</sub> |
|       |                | A5           | 5CG1_A | Homo sapiens | SM      |
|       |                | A5           | 5RV_A | Homo sapiens | GO<sub>4</sub>, PTR<sub>4</sub> |
|       |                | A5           | 40G3_A | Homo sapiens | XG      |
|       |                | A5           | 3X2Z_A | Homo sapiens | FLL, TPO<sub>4</sub> |
|       |                | A5           | 1LY7_A | Homo sapiens | ADP, MG, TPO<sub>4</sub> |
|       |                | A5           | 5NNR_A | Homo sapiens | ATP, MG, SO4, TPO<sub>4</sub> |
|       |                | A5           | 5D1_A | Homo sapiens | ATP, MG, SO, TPO<sub>4</sub> |
|       |                | A5           | 4DVE_A | Homo sapiens | ADP, MG |

Results and discussion:
The structures retained by the filters used during the study of the structural differentiation between P-kinase, before and after phosphorylation of the activation loop, are listed in Table 1. The interaction tables can be found in the supplementary material. Cyclin-dependent kinase 2 (CDK2) [34] is an essential component of the cell cycle machinery, with maximal activity during S phase. As its name indicates, its functionality depends on the presence of cyclin [35]. Phosphorylation at T14 or Y15 residues of CDK2 causes inactivation of CDK2, whereas phosphorylation at T160 increases its activity [36-37]. The importance of the phosphorylation of the latter was confirmed during its mutation to alanine, leading to a decrease of the CDK2 activity five times [38]. The phosphorylation of CDK2 at T160 leads this amino acid to interact by hydrogen...
Glycogen synthase kinase three beta (GSK3β) is a protein of the GSK family of CMGC group [40], phosphorylates glycogen synthase to inactivate, participate in the Wnt signaling [41], which is involved in energy metabolism and neuronal cell development [41]. Phosphorylation of GSK3 at Y216 position acts as an activator of the β-ACTIN, participating in the Wnt signaling family of CMGC group [2063 (online)].

Phosphorylation of GSK3 at Y216 position acts as an activator of the β-ACTIN, participating in the Wnt signaling family of CMGC group [2063 (online)].

Phosphorylation of GSK3 at Y216 position acts as an activator of the β-ACTIN, participating in the Wnt signaling family of CMGC group [2063 (online)].

Phosphorylation of GSK3 at Y216 position acts as an activator of the β-ACTIN, participating in the Wnt signaling family of CMGC group [2063 (online)].

Phosphorylation of GSK3 at Y216 position acts as an activator of the β-ACTIN, participating in the Wnt signaling family of CMGC group [2063 (online)].

Phosphorylation of GSK3 at Y216 position acts as an activator of the β-ACTIN, participating in the Wnt signaling family of CMGC group [2063 (online)].
segment [56]. Moreover, SYK kinase dual phosphorylation does not lead to a significant change in interactions and does not affect specific sub-domains, except in the case of the disappearance of two DS12 and G514 hydrogen interactions of the motif DFG with N381 and F382. This last amino acid is marked by the most significant deviation at the dihedral angle (φ) = 167.4°. The low number of interactions may be due to the stability of the activation loop in the two cases of phosphorylation, since the secondary structure of the unphosphorylated form has the two sheets (β6-β9 and β10-β11). Single phosphorylation does not result in a large propagation of this effect all along the activation loop, because this propagation only affects the C-terminal two amino acids in the case of single phosphorylation (T530, H531). Phosphorylation of PAK1 kinase does not result in a fluctuation transfer at the activation loop, which results in a small number of interactions that do not affect specific subdomains. The activation loop is stabilized with the beginning of the formation of the β10-β11 sheet, and the β6-β9 sheet.

Aurora is a member of the AUR kinase family [57]. The kinase is located on mitotic centrosomes and microtubules, required for centrosome maturation [58]. Activation of Aurora requires binding of the TPX2 complex [59], but is enhanced by phosphorylation at position T287 and is suppressed when position T288 is also phosphorylated [60]. In case of AURORA kinase monophosphorylation at T287, the activation loop remains stabilized with the beginning of the formation of the β10-β11 sheet and the β6-β9 sheet. This stability is reflected in the small fluctuation of the entire segment. All these fluctuations do not prevent some structural changes scattered between the two lobes and marked by the conservation of multiple interactions between Q177 and W277, and the disappearance of a hydrogen interaction with G276, which belongs to the DFG motif. As regards monophosphorylation at T288, in an earlier work by Bayliss, we have found a difference in the analyzed structure, since it is complexed with TPX2. We note that we do not have the same interactions with T288, whereas there is the appearance of interactions that resemble the case of T287 phosphorylation. This finding allows us to say in the case of the work of Bayliss, that these interactions are not due to the phosphorylation but to the complexation of TPX2 [1]. For the doubly phosphorylated structure, it is noted that R285, of the phosphorylated structure T288, binds with R180 and R255, the same goes with those where the phosphorylated structure binds to T287 as if there exists some competition between the T287 and T288 phosphorylated structures.

**Conclusion:**

Results show that the structural adaptation of the activation loop after phosphorylation is mainly due to hydrogen bonds formed between the phosphate group with amino groups (R-NH2) of lysines or with guanidine groups (R-Ch3N+) of arginines. The multiplicity of these phosphate groups represents anchors, which stabilize the activation loop. The strongest anchor belonging to the N-terminal concerns arginines or lysines of the Ca and β9 and also arginines of the HRD motif, as well. The second remarkable anchor, which belongs to C terminal, concerns arginines or lysines of the βEF and leads to a rearrangement of the loop P+1 amino acids. Moreover, these anchors allow interactions propagation towards the N-terminal lobe of the activation segment. Further, it should be noticed that the activation loop stability is conditioned by the existence of interactions between the main chains of the activation loop and the αEF/αF loop. Besides, the structures containing β6-β9 sheets in the N-terminal or β10-β11 in the middle of the activation loop strengthen its stability during phosphorylation. Finally, we find out that interactions’ variations are acting on the most essential regions at the level of kinases with the hydrogen bonds, affecting the most conserved motifs in kinases: DFG for the SYK kinase, APE for the PAK1 kinase and HRD for CDK2, MAPK14, PDK1, PAK1, Aurora as well as other ones which react on the important loops, like GSK3β on Mg2+ loop and MAPK14 on P+1 loop.

**Acknowledgments:**

We want to thank Mr. R. Boulguid for support.

**Author contributions:**

Adil Ahiri did data analysis and interpretation, article writing. Aziz Aboulmouhajir did data analysis and interpretation, article reviewing, article writing; Crtomir Podlipnik did data analysis and interpretation, article reviewing and Hocine Garmes did article reviewing. The open access charges for this article is fully sponsored by Biomedical Informatics (P) Ltd, India.

**Conflicts of Interest:**

The authors declare that they have no conflict of interest.

**References:**

[1] Bayliss R et al. Open Biology 2012 2:120136. [PMID: 23226601]
[2] Johnson LN et al. FEBS Letters 1998 430:1. [PMID: 9678585]
[3] Nolen B et al. Molecular Cell 2004 15:661. [PMID: 15350212]
[4] Luhovy AX et al. J Biol Chem. 2012 285:5446. [PMID: 22203681]
[5] Tokmakov AA et al. Biochim Biophys Acta. 2000 1476:66. [PMID: 10606768]
[6] Xiong S et al. Structure. 2018 26:1101. [PMID: 29983373]
[7] Donella-Deana A et al., Biochemistry. 1998 37:1438. [PMID: 9477973]
[8] Cybulsky AV et al. PLoSOne. 2017 12:e0177226. [PMID:28475647]
[9] Panneton V et al. J Biol Chem. 2015 290:20960. [PMID: 26170449]
[10] Pike AC et al. EMBO J. 2008 27:704. [PMID: 18239682]
[11] Wybenga-Groot LE et al. Cell Signal. 2014 26:1825. [PMID: 24815189]
[12] Oliver AW et al. Trends Biochem Sci. 2007 32:351. [PMID: 17627826]
[13] Adams JA. Biochemistry. 2003 42:601. [PMID:12534271]
[14] Bhullar KS et al. Molecular cancer. 2018 17:48. [PMID:29455673]
[15] Duong-Ly KC & Peterson JR. Current Protocols in Pharmacology. 2013 6:2. [PMID: 23456613]
[16] Quintale SB & Orchard S. Molecular & Cellular Proteomics. 2008 7:1409. [PMID: 18436524]
[17] Manning G et al. Science. 2002 298:1912. [PMID: 12471243]
[18] Wu CH et al. Nucleic Acids Research. 2006 34:D1871. [PMID: 16381842]
[19] Berman HM et al. Acta Crystallogr D. Biol Crystallogr. 2006 2000:235. [PMID: 12037327]
[20] Allschnel SF et al. Nucleic Acids Res. 1997 25:3389. [PMID: 9254694]
[21] Hornbeck PV et al. Proteomics 2004 4:1551. [PMID: 15174125]
[22] Pettersen EF et al. Journal of Computational Chemistry. 2004 25:1605. [PMID: 15264254]
[23] Li X et al. Analytical Chemistry. 2016 88:11468. [PMID: 27808495]
[24] Kooistra AJ et al. Nucleic Acids Research. 2015 44:D365. [PMID: 26496499]
[25] Grant BJ et al. Bioinformatics. 2006 22:2695. [PMID: 16940322]
[26] Tsang E et al. Journal of Biological Chemistry. 2008 283:326509. [PMID: 18818202]
[27] Sigrist CJ et al. Nucleic Acids Research. 2012 41:D344. [PMID: 23161676]
[28] Fiser A & Sali A. Methods in Enzymology. 2003 374:461. [PMID: 14696385]
[29] Scouras AD & Daggett V. Protein Science. 2011 20:34. [PMID: 21280126]
[30] Gaillard T et al. Journal of Chemical Information and Modeling. 2013 53:2471. [PMID: 23957210]
[31] Sellers BD et al. Proteins: Structure, Function, and Bioinformatics. 2010 78:2490. [PMID: 20620354]
[32] Zhang T et al. 2010; Proteins: Structure, Function, and Bioinformatics. 78:3353. [PMID: 20818661]
[33] Das AA & Krishna R. Bioinformation. 2016 12:78. [PMID: 28104965]
[34] Harper JW & Adams P. Chemical Review. 2001 101:2511. [PMID: 11749386]
[35] van den Heuvel S & Harlow E. Science. 1993 262:2050. [PMID: 8266103]
[36] Bartova I et al. Protein Science. 2004 13:1449. [PMID: 15131614]
[37] Bartova I et al. Protein Science. 2008 17:22. [PMID: 18042686]
[38] Connell-Crowley L et al. Molecular Biology of the Cell. 1993 4:79. [PMID: 8443411]
[39] Johnson LN. Biochemical Society Transactions. 2009 37:627. [PMID: 19614568]
[40] Cohen P & Frame S Nature Reviews Molecular Cell Biology. 2001 2:769. [PMID: 11584308]
[41] Wu D & Fan W. Trends in Biochemical Science. 2010 35:161. [PMID: 19884009]
[42] Dajani R et al. Cell. 2001 105:721. [PMID: 11440715]
[43] Krupa A et al. Journal of Molecular Biology. 2004 339:1025. [PMID: 15178245]
[44] Varjosalo, M., et al. Cell reports. 2013 3:1306. [PMID: 23602568]
[45] Munshi A., & Ramesh, R. Genes & Cancer. 2013 4:401. [PMID: 24349638]
[46] Zhang Y et al. Journal of Biological Chemistry. 2008 283:26591. [PMID: 18669639]
[47] Beesnstock J. Trends in Biochemical Sciences. 2016 41:938. [PMID: 27594179]
[48] Brazil DP & Hemmings BA. Trends in biochemical science. 2001 26:657. [PMID: 11701324]
[49] Lemmon MA & Ferguson KM. Biochemical Journal. 2000 350:1. [PMID: 1092682]
[50] Frödin M & Gammeltoft S. Molecular and Cellular Endocrinology. 1999 151:65. [PMID: 10411321]
[51] Komander D et al. Journal of Biological Chemistry. 2005 280:18797. [PMID: 15741170]
[52] Kurosaki T et al. The Journal of experimental medicine. 1994 179:1725. [PMID: 7513017]
[53] Cheng AM et al. Nature. 1995 378:303. [PMID: 7477353]
[54] Aoki Y. et al. Proceedings of the National Academy of Science. 1994 91:10606. [PMID: 7524098]
[55] Kurosaki T. et al. Journal of Experimental Medicine. 1995 182:1815. [PMID: 7500027]
[56] Telesco SE et al. Biotechnology Journal. 2013 8:1452. [PMID: 24318637]
[57] Keen N & Taylor S. Nature Reviews Cancer. 2004 4:927. [PMID: 15573114]
[58] Berdnik D & Knoblich JA. Current Biology. 2002 12:640. [PMID: 11967150]
[59] Dodson CA et al. Biochemical Society Transactions. 2013 41:1037. [PMID: 23863175]
Rowan FC et al. ACS Chemical Biology. 2013 8:2184. [PMID: 23924325]

License statement: This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License.
Analysis of structural variations of the unphosphorylated (PDB ID: 1FIN-A) and T160 phosphorylated (PDB ID: 4EOQ-A) structures of the CDK2 kinase. The localization of the amino acid and its kind of interaction are as follows:

Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction/ green: a hydrophobic interaction / violet: anion-π interaction/ white: multiple interactions.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|---------------------|----------------------|-----------------------|
|                     | Cartesian analysis    | Dihedral analysis      | RMSF<sub>Δphi</sub> | RMSF<sub>Δpsi</sub> | UNPHOSPHO | PHOSPHO |
| F152                | 0.5 0.2 8.9 68.7     |                       |                      |                      |           |         |
| G153                | 1.4 0 18.9 130.3     |                       |                      |                      |           |         |
| T160                | 2.7 2.5 25.7 17.8    |                       |                      |                      |           |         |
| Y156                | 3.5 3.4 7.0 68.0     |                       |                      |                      |           |         |
| R157                | 2.8 2.6 22.3 134.7   |                       |                      |                      |           |         |
| T158                | 2.7 2.7 21.7 15.7    |                       |                      |                      |           |         |
| Y159                | 3.5 3.2 12.1 14.5    |                       |                      |                      |           |         |
| T160 *               | 3.3 3.7 6.7 79.9     |                       |                      |                      |           |         |
| H161                | 3.4 3.7 10.4 2.6     |                       |                      |                      |           |         |

Analysis of structural variations of the unphosphorylated (PDB ID: 2O5K-A) and Y216 phosphorylated (PDB ID: 2OW3-A) structures of the GSK3β kinase. The localization of the amino acid and its kind of interaction are as follows:

Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction/ violet: anion-π interaction/ white: multiple interactions.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|---------------------|----------------------|-----------------------|
|                     | Cartesian analysis    | Dihedral analysis      | RMSF<sub>Δphi</sub> | RMSF<sub>Δpsi</sub> | UNPHOSPHO | PHOSPHO |
| Y216                | 1.5 0.5 25.0 14.0    |                       |                      |                      |           |         |
| T252                | 0.5 0.1 15.0 10.5    |                       |                      |                      |           |         |
| G259                | 0.6 0.5 6.0 3.5      |                       |                      |                      |           |         |
| N285                | 0.9 0.5 13.0 14.0    |                       |                      |                      |           |         |

©Biomedical Informatics (2019)
### Cartesian analysis

| AMINO ACID POSITION | DEVIATION | RMSD<sub>x</sub> | RMSD<sub>y</sub> | DEVIATION | RMSD<sub>x</sub> | RMSD<sub>y</sub> |
|---------------------|-----------|----------------|----------------|-----------|----------------|----------------|
| S234                | 2.5       | 3.1            | 6.0            | 2.0       | 5.0            | 1.0            |
| K235                | 3.0       | 3.4            | 7.9            | 2.5       | 2.9            | 3.0            |
| Q236                | 2.5       | 4.3            | 3.8            | 3.5       | 4.3            | 3.8            |
| A237                | 2.8       | 3.7            | 5.6            | 9.8       | 2.0            | 3.0            |
| R238                | 2.8       | 0.5            | 3.0            | 3.1       | 2.0            | 3.0            |
| A239                | 2.4       | 2.1            | 2.1            | 2.0       | 2.0            | 2.0            |
| E240                | 3.1       | 4.3            | 4.5            | 8.0       | 3.0            | 4.5            |
| K241                | 3.6       | 4.8            | 4.3            | 14.8      | 1.2            | 3.0            |
| F242                | 2.6       | 2.1            | 12.1           | 7.3       | 2.0            | 3.0            |

### Dihedral analysis

| AMINO ACID POSITION | DEVIATION | RMSD<sub>x</sub> | RMSD<sub>y</sub> | DEVIATION | RMSD<sub>x</sub> | RMSD<sub>y</sub> |
|---------------------|-----------|----------------|----------------|-----------|----------------|----------------|
| S234                | 2.5       | 3.1            | 6.0            | 2.0       | 5.0            | 1.0            |
| K235                | 3.0       | 3.4            | 7.9            | 2.5       | 2.9            | 3.0            |
| Q236                | 2.5       | 4.3            | 3.8            | 3.5       | 4.3            | 3.8            |
| A237                | 2.8       | 3.7            | 5.6            | 9.8       | 2.0            | 3.0            |
| R238                | 2.8       | 0.5            | 3.0            | 3.1       | 2.0            | 3.0            |
| A239                | 2.4       | 2.1            | 2.1            | 2.0       | 2.0            | 2.0            |
| E240                | 3.1       | 4.3            | 4.5            | 8.0       | 3.0            | 4.5            |
| K241                | 3.6       | 4.8            | 4.3            | 14.8      | 1.2            | 3.0            |
| F242                | 2.6       | 2.1            | 12.1           | 7.3       | 2.0            | 3.0            |

Analysis of structural variations of the unphosphorylated (PDB ID: 2BIY) and 1P38 phosphorylated (PDB ID: 3P3Y) structures of MAPK14 kinase. The localization of the amino acid and its kind of interaction are as follows: AMINO ACID IN BOLD LINE BELONG TO THE ACTIVATION SEGMENT; Yellow: a hydrogen bond interaction; green: a hydrophobic interaction; white: multiple interactions.

### AMINO ACID POSITION

| AMINO ACID POSITION | Structural variation | Interaction variation |
|---------------------|----------------------|----------------------|
| S234                | 2.5                  | 3.1                  |
| K235                | 3.0                  | 3.4                  |
| Q236                | 2.5                  | 4.3                  |
| A237                | 2.8                  | 3.7                  |
| R238                | 2.8                  | 0.5                  |
| A239                | 2.4                  | 2.1                  |
| E240                | 3.1                  | 4.3                  |
| K241                | 3.6                  | 4.8                  |
| F242                | 2.6                  | 2.1                  |

Analysis of structural variations of the unphosphorylated (PDB ID: 2BIY) and 1P38 phosphorylated (PDB ID: 3P3Y) structures of MAPK14 kinase. The localization of the amino acid and its kind of interaction are as follows: AMINO ACID IN BOLD LINE BELONG TO THE ACTIVATION SEGMENT; Yellow: a hydrogen bond interaction; green: a hydrophobic interaction; white: multiple interactions.

### AMINO ACID POSITION

| AMINO ACID POSITION | Structural variation | Interaction variation |
|---------------------|----------------------|----------------------|
| S379                | 2.4                  | 3.0                  |
| G380                | 2.4                  | 4.2                  |
| N381                | 3.9                  | 4.2                  |
| F382                | 2.9                  | 4.2                  |
| G383                | 2.5                  | 4.2                  |

Analysis of structural variations of the unphosphorylated (PDB ID: 1P38) and 1P38 phosphorylated (PDB ID: 3P3Y) structures of MAPK14 kinase. The localization of the amino acid and its kind of interaction are as follows: AMINO ACID IN BOLD LINE BELONG TO THE ACTIVATION SEGMENT; Yellow: a hydrogen bond interaction; green: a hydrophobic interaction; white: multiple interactions.
Analysis of structural variations of the unphosphorylated (PDB ID: 1YHW) and Y525 and Y526 phosphorylated (PDB ID: 3SRV) structures of the Syk kinase. The localization of the amino acid and its kind of interaction are as follows: Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction; Green: a hydrophobic interaction; Pink: a cation-π interaction; Orange: lone pair - π interaction; White: multiple interactions.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|----------------------|----------------------|-----------------------|
|                      | Cartesian analysis    | Dihedral analysis     | RMSF<sub>sc</sub> | RMSF<sub>bb</sub> | σφ | σψ |
|                      |                      |                       | UNPHOSPHO | PHOSPHO |
| G390                 | 1.4                  | 0.0                   | 42.3      | 3.4      |     |     |
| N351                 | 0.0                  | 0.0                   | 0.9      | 1.2      |     |     |
| F382                 | 2.1                  | 3.4                   | 31.3      | 107.3    | K333 |     |
| G383                 | 1.2                  | 0.0                   | 48.9      | 27.9      |     |     |
| K390                 | 0.3                  | 0.0                   | 11.4      | 35.9      |     |     |
| K409                 | 0.7                  | 0.0                   | 19.7      | 165.7     |     |     |
| H408                 | 2.0                  | 3.5                   | 64.0      | 32.5      |     |     |
| E407                 | 2.3                  | 3.8                   | 50.1      | 135.7     |     |     |
| A406                 | 1.8                  | 2.2                   | 45.2      | 8.9       |     |     |
| I410                 | 2.1                  | 3.1                   | 8.2       | 99.5      |     |     |
| P411                 | 1.8                  | 3.2                   | 5.3       | 83.3      |     |     |
| A442                 | 0.5                  | 2.3                   | 59.0      | 133.7     |     |     |
| A441                 | 0.6                  | 1.2                   | 14.4      | 34.3      |     |     |
| E442                 | 0.5                  | 1.8                   | 64.0      | 47.9      |     |     |
| S443                 | 0.1                  | 0.3                   | 55.9      | 2.8       |     |     |
| Y525*                | 0.1                  | 0.7                   | 0.3      | 1.7       | K348 | K548 |
| Y526*                | 0.1                  | 0.4                   | 1.7      | 1.8       |     |     |
| T580                 | 1.5                  | 2.7                   | 34.6      | 25.2      |     |     |
| H531                 | 1.3                  | 2.9                   | 10.9      | 4.8      | 1944 | Y547 |

Analysis of structural variations of the unphosphorylated (PDB ID: 4XG3) and Y525 and Y526 phosphorylated (PDB ID: 4YJT) structures of the PAK1 kinase. The localization of the amino acid and its kind of interaction are as follows: Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction; Pink: a cation-π interaction; White: multiple interactions.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|----------------------|----------------------|-----------------------|
|                      | Cartesian analysis    | Dihedral analysis     | RMSF<sub>sc</sub> | RMSF<sub>bb</sub> | σφ | σψ |
|                      |                      |                       | UNPHOSPHO | PHOSPHO |
| I276                 | 0.9                  | 1.2                   | 3.9      | 51.0      |     |     |
| G277                 | 0.3                  | 0.0                   | 96.4      | 3.0       |     |     |
| Q278                 | 0.4                  | 0.2                   | 3.2      | 52.3      |     |     |
| T279                 | 0.6                  | 0.0                   | 49.4      | 4.4       |     |     |
| F423*                | 0.1                  | 0.1                   | 3.1      | 4.4       | Y441 | Y433 |
| T437                 | 0.7                  | 0.4                   | 8.5      | 117.8     | E433 | E434 |
| R438                 | 0.4                  | 0.6                   | 132.8     | 3.8       |     |     |
| K439                 | 0.8                  | 0.1                   | 3.6      | 97.1      |     |     |
| A440                 | 0.3                  | 0.4                   | 105.4     | 1.1       |     |     |

Analysis of structural variations of the unphosphorylated (PDB ID: 1YHW) and T423 phosphorylated (PDB ID: 3FXZ) structures of the PAK1 kinase. The localization of the amino acid and its kind of interaction are as follows: Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|----------------------|----------------------|-----------------------|
|                      | Cartesian analysis    | Dihedral analysis     | RMSF<sub>sc</sub> | RMSF<sub>bb</sub> | σφ | σψ |
|                      |                      |                       | UNPHOSPHO | PHOSPHO |
| H176                 | 0.6                  | 0.4                   | 5.7      | 96.0      |     |     |
| Q277                 | 0.4                  | 0.6                   | 80.4      | 9.7       | G296 | W297 |
Analysis of structural variations of the unphosphorylated (PDB ID: 4DEE-A) and T287 phosphorylated (PDB ID: 5DT3-A) structures of the AURORA kinase. The localization of the amino acid and its kind of interaction are as follows: Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction/pi$^2$ interaction/cation-π interaction/white: multiple interactions.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|---------------------|----------------------|-----------------------|
|                     | Cartesian analysis    | Dihedral analysis     | RMSF$_{bb}$ | RMSF$_{sc}$ | σ($\phi$) | σ($\psi$) | UNPHOSPHO | PHOSPHO |
| S284                | 0.3                  | 0.9                  | 1.2         | 13.9       | R255      | M305     |
| R285                | 1.2                  | 4.9                  | 21.9        | 21.2       | ----      | R255     | M305     |
| R286                | 2.0                  | 4.7                  | 6.6         | 10.8       | H176      | ----      |
| T287                | 2.4                  | 3.1                  | 4.7         | 30.2       | ----      | ----      |
| T288 *              | 1.9                  | 3.5                  | 22.0        | 7.8        | ----      | R285     |
| L289                | 1.5                  | 3.1                  | 7.0         | 11.2       | ----      | ----      |
| C290                | 1.6                  | 3.6                  | 22.2        | 28.6       | K143      | V354     |

Analysis of structural variations of the unphosphorylated (PDB ID: 4DEE-A) and T287 phosphorylated (PDB ID: 5DNR-A) structures of the AURORA kinase. The localization of the amino acid and its kind of interaction are as follows: Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction/green: a hydrophobic interaction/pi$^2$ interaction/cation-π interaction/white: multiple interactions.

| AMINO ACID POSITION | STRUCTURAL VARIATION | INTERACTION VARIATION |
|---------------------|----------------------|-----------------------|
|                     | Cartesian analysis    | Dihedral analysis     | RMSF$_{bb}$ | RMSF$_{sc}$ | σ($\phi$) | σ($\psi$) | UNPHOSPHO | PHOSPHO |
| H280                | 1.1                  | 1.2                  | 2.2         | 30.2       | R255      | V252     |
| A281                | 1.1                  | 2.4                  | 48.0        | 20.9       | R251      | R251     |
| R285                | 0.6                  | 0.9                  | 16.2        | 41.4       | ----      | ----      |
| R286                | 0.7                  | 1.8                  | 17.9        | 28.5       | H176      | L287     |
| T287 *              | 1.4                  | 3.5                  | 106.1       | 142.6      | ----      | ----      |
| T288 *              | 0.5                  | 3.9                  | 22.2        | 37.6       | ----      | ----      |
| L289                | 0.3                  | 0.2                  | 56.1        | 10.3       | ----      | V354     |
| C290                | 0.9                  | 0.1                  | 11.3        | 130.3      | K143      | ----      |
| G291                | 0.6                  | 0.0                  | 55.3        | 38.5       | ----      | ----      |

Analysis of structural variations of the unphosphorylated (PDB ID: 4DEE-A) and T287 and T288 phosphorylated (PDB ID: 1OL7) structures of the AURORA kinase. The localization of the amino acid and its kind of interaction are as follows: Amino acids in bold line belong to the activation segment; Yellow: a hydrogen bond interaction/green: a hydrophobic interaction/pink: a cation-π interaction/violet: anion-π interaction/orange: lone-pair - π interaction.
