Acta Biochimica Polonica

Vol. 44 No. 3/1997

557–564

QUARTERLY

Modelling of active forms of protein kinases: p38 — a case study*

Krzysztof Ginalski[1, 2], Bogdan Lesyng[1, 2,✉], Janusz Sowadski[3] and Michał Wojciechowski[1, 2]

[1]*Interdisciplinary Centre for Mathematical and Computational Modelling, Warsaw University, A. Pawińskiego 5A, 02-106 Warsaw, Poland;*
[2]*Department of Biophysics, Warsaw University, F. Żwirki i S. Wigury 93, 02-089 Warsaw, Poland;*
[3]*Department of Biochemistry, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 20111, U.S.A.*

Key words: mitogen activated protein kinase, p38 kinase, protein kinase A, extracellular signal-regulated kinase 2, phosphorylation, molecular modelling, molecular mechanics, molecular dynamics, drug design

An active form of p38 protein kinase, belonging to the mitogen-activated protein kinases subfamily, has been designed based on crystallographically known structures of two other kinases, an active form of protein kinase A (PKA) and an inactive form of extracellular signal-regulated kinase 2 (ERK2). The modelling procedure is described. Its general scheme can also be applied to other kinases. The structure of the active forms of p38 and PKA is very similar in the region which binds the substrate. The ATP-binding mode is very similar in the active forms of all the three studied kinases. Models of the active forms allow for further studies on transphosphorylation processes at the molecular level, and modelling of inhibitors competitive with ATP and/or substrates.

The family of protein kinases is large and diverse. The enzymes differ in size and subcellular distribution. Protein kinases and phosphatases play a key role in signalling and posttranslational control of biochemical processes (for review see e.g. [1]). The best known protein kinase is the cyclic AMP-dependent protein kinase (cAPK), called also protein kinase A (PKA). Crystalline structure of the enzyme in complex with a protein kinase inhibitor (PKI) was solved first [2, 3]. Several other forms of PKA complexed with the inhibitor and/or MnATP have been crystallised and their structures determined [4, 5]. The structures of several other protein kinases have also been reported, among them: cyclin-dependent kinase 2 (CDK2) [6], extracellular signal-regulated kinase 2 (ERK2) [7], twichin kinase [8], and insulin receptor kinase (IRK) [9]. Although the sequence identity among many protein kinases is small, their secondary structures reveal several common features [10]. All known protein kinases contain a two-domain conserved

*The studies were supported by the State Committee for Scientific Research (KBN 8T11F 006 09).
✉Corresponding author: e-mail: lesyng@icm.edu.pl.; phone: (+22)+879-91-00; fax: (+22)+874-91-15.

Abbreviations: cAPK, cAMP-dependent protein kinase; CDK, cyclin-dependent kinase; ERK, extracellular signal-regulated kinase; IL, interleukin; IRK, insulin receptor kinase; MAPK, mitogen-activated protein kinase; MKK, mitogen-activated protein kinase kinase; MLCK, myosin light chain kinase; BPDB, Brookhaven Protein Data Bank; PKA, protein kinase A; PKI, protein kinase inhibitor; r.m.s., root mean square.

catalytic core, consisting of about 260 residues. The amino-terminal domain, the "upper" one, consists of a twisted β-sheet, while the carboxy-terminal domain, the "lower" one, is composed almost entirely of α-helices. The domains can form either an "open" or "closed" form [4, 11]. The first is capable of accommodating ATP and a substrate. After the ATP and substrate have been docked, the domain can close and the enzyme is thus transformed, ready for transphosphorylation process. Figure 1 shows traces of the α-carbon atoms in the closed and open forms of PKA. A detailed analysis of the crystallographic data shows that the internal structure of the domains in the closed and open forms is very similar. This observation was taken advantage of in the modelling procedure of the active forms (see below). The knowledge of the enzyme active forms is required for detailed, quantum-dynamics studies (see e.g. [12]) of the phosphorylation mechanisms, as well as for rational design of inhibitors (potential drugs).

There are nine residues which are conserved in their 3D positions in all known structures of protein kinases. They play a crucial role in the binding of ATP and substrates, and are responsible for the transphosphorylation processes. ATP, which is the γ-phosphate donor, binds in a pocket between the above mentioned domains. The substrate binding site is located on the outer part of the protein, covering the lips of the cleft between the domains. In PKI there are several basic residues preceding, and one hydrophobic residue following the phosphorylation site (Ser/Thr). The preferred sequence of the PKA substrates is: Arg-Arg-X-Ser/Thr-Y, where X is any residue and Y a hydrophobic one.

The purpose of this study was to develop an effective method for designing of active forms of biologically important protein kinases whose crystallographic structures are not

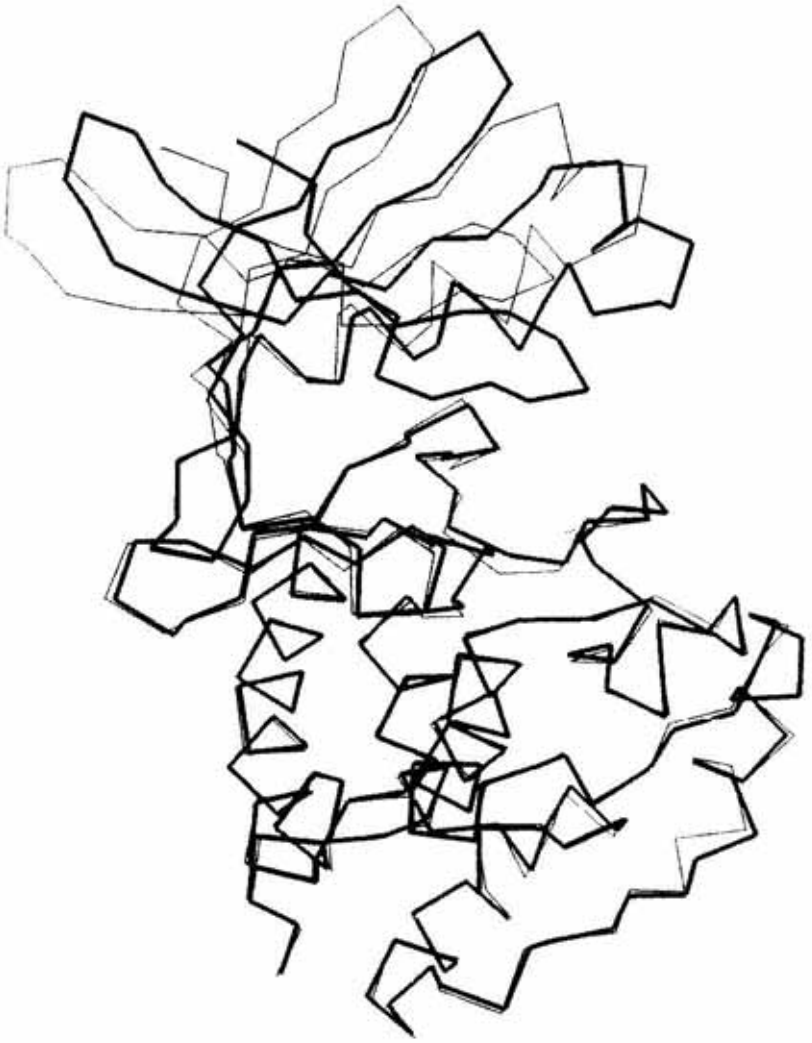

Figure 1. The α-carbon trace diagram of the closed (bold) and open (thin) conformations of the PKA catalytic core.

The carboxy-terminal (lower) domains are superimposed, showing the rotation of the upper domain in the open conformation.

known. For this study p38 kinase has been chosen. Its active form is compared with the known active form of PKA, their common features are analysed and will be used for further analysis of the transphosphorylation processes catalysed by kinases.

Kinase p38 belongs to the MAPK (mitogen-activated protein kinases) subfamily. MAPK cascades represent one of the major signal systems used by eukaryotic cells to transduce extracellular signals into cellular responses. Kinases MKK3, MKK4 and MKK6 are nearly equally active when phosphorylating p38 [13, 14]. Kinases of the MAPK subfamily achieve maximum activity only when both Tyr and Thr, located in the activation loop, are phosphorylated. These are Thr-183 and Tyr-185 in the case of extracellular signal-regulated kinase 2 (ERK2), and Thr-180 and Tyr-182 in the case of p38 (see Fig. 2 for the alignment scheme of PKA, ERK2 and p38). p38 was first identified in studies on a complex glycolipid found in the outer membrane of Gram-negative bacteria. Further studies have shown that p38 activation can also be observed in cells treated with cell wall components from Gram-positive bacteria, proinflammatory cytokines such as tumour necrosis factor-α (TNF-α) and interleukin-1 (IL-1), physical stress such as high osmolarity, as well as chemical stress such as exposition to H_2O_2 [15, 16]. Activation of p38 may be essential for some cellular responses associated with acute or chronic inflammation.

ERK2 also belongs to the MAPK subfamily. Its inactive 3D structure is known [7]. As mentioned above, the active form of PKA is also known. Neither the active nor inactive form of p38 has been determined. The first two structures will be used in this study as the reference systems for constructing the active form of p38.

Both in p38 and ERK2 the conserved core has an insertion between β4 and β5 strands as compared with PKA (see Fig. 2). p38 has also one residue more than ERK2 in the loop between β5 strand and the αD helix, and it's length is the same as the length of the loop in PKA. In p38 the helix αB, present in PKA, doesn't exist. In the lower domain, p38 as compard to PKA has an insertion after the helix αG, present also in ERK2. Most important, the activation loop in ERK2 is longer by six residues than the loops of PKA and p38. All these features were taken into account in the procedure of p38 modelling.

MODELLING SCHEME

Taking into consideration that all crystallographically known protein kinases bind ATP in the same manner, i.e. in the cleft between the "upper" (small) and "lower" (large) domains, it can be concluded that the relative orientation of these domains in their active forms should be similar. Consequently, the PKA active form can be treated as a template for relative orientations of the small and large domains in the active forms of other kinases. The modelling procedure described below includes some of the techniques previously applied in modelling of other kinases, in particular myosin light chain kinase (MLCK) [17] and EGF receptor [18].

The model of the p38 active form was designed using the Homology program [19] based on the inactive ERK2 crystallographic structure [7]. A two-step procedure was applied. First the structure of the ERK2 active form was built based on the crystal structure of its inactive form. Next the p38 active form was obtained by replacing side chains of the ERK2($Mg^{2+}{}_2ATP^{-4}$) model with the p38 side chains according to the alignment scheme depicted in Fig. 2.

The model of the ERK2($Mg^{2+}{}_2ATP^{-4}$) active form was designed based on the crystal structure of the PKA active form and the ERK2 inactive form. All atomic positions were taken from the ERK2 crystallographic structure and those of the lower lobe were unchanged. The upper lobe was treated as a rigid body and transformed to obtain its relative orientation with respect to the lower lobe the same as in the PKA active form. The linker between the domains (Asp104-Asp109) was adjusted manually. Finally $Mg^{2+}{}_2ATP^{-4}$ was docked in the same orientation as in the PKA crystallographic structure. Refinement of the structure was done using the energy minimization procedure as described below for the p38 model. The en-

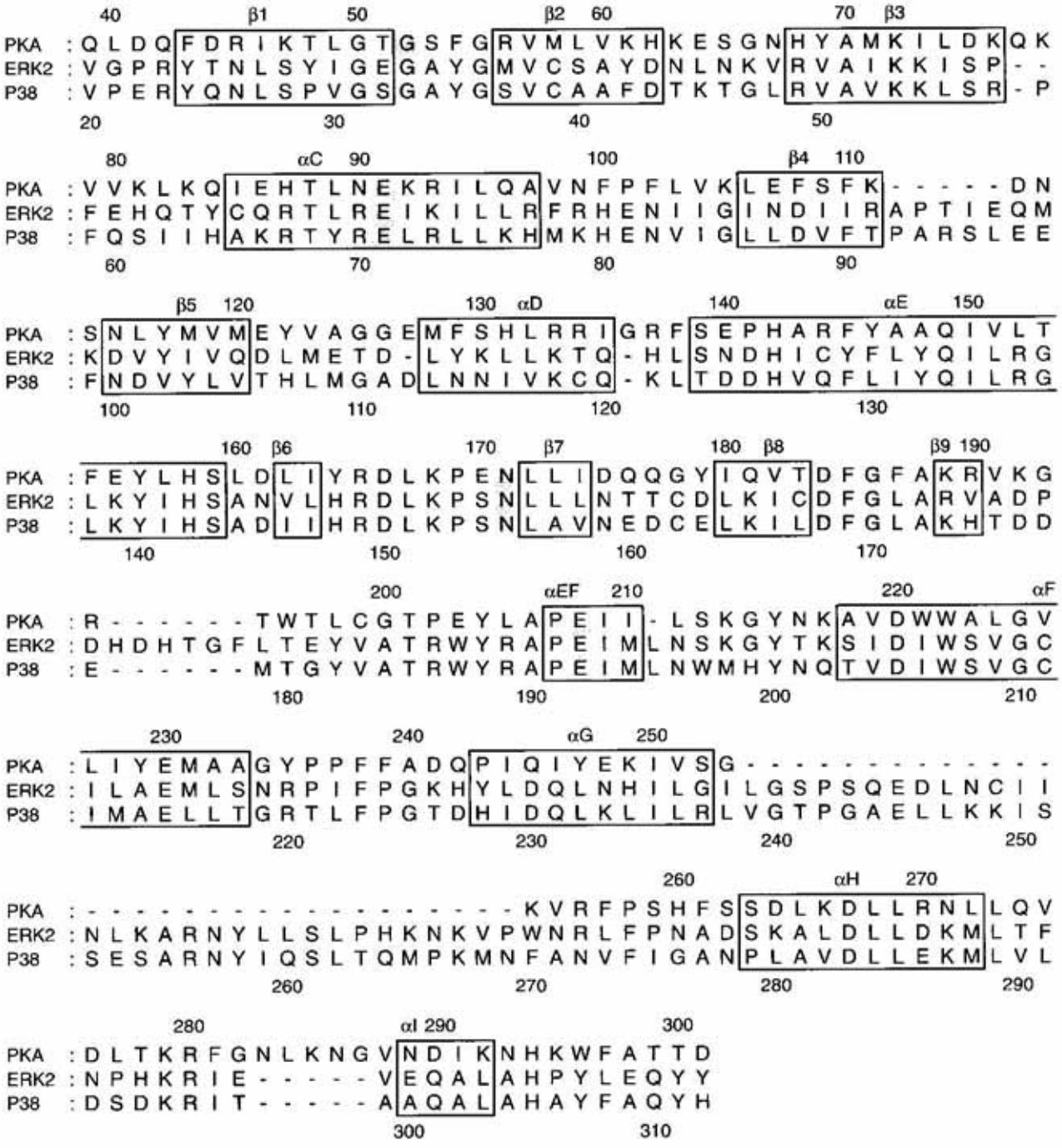

Figure 2. Sequence alignment of the PKA, ERK2 and p38 catalytic cores.

Primary sequences were aligned by using invariant residues in the crystal strcucture models of PKA and ERK2. Boxed sequences indicate α-helices and β-strands of PKA that are conserved in ERK2 and p38. Highly conserved sequences are shaded.

ergy optimizations were performed with the AMBER force field [20] applying the steepest descent and conjugate gradient methods.

Regarding the procedure of p38 modelling, positions of the ERK2($Mg^{2+}{}_2ATP^{-4}$) backbone atoms in the conserved regions and

loops of the same length were left unchanged. The p38 side chains attached to the conserved regions followed the ERK2 side chain positions wherever possible. The loops with insertions in p38 were modelled using the information contained in the Brookhaven Protein Data Bank (BPDB). We assumed that conformation of an insertion is determined by its two neighbouring regions. These regions, separated by a gap of a length equal to the length of the insertion, were searched in the bank. The detected structural fragments were adopted for modelling of the insertion by assuming that the backbone(s) of the gap region(s) and the backbone(s) of the required insertion(s) are identical. In the case when few candidates for such an insertion were available, this was chosen which resulted in the best fitting of the insertion into the existing structure of p38. Regarding the deletion in the activation loop, it was generated manually. In order to eliminate tensions, the model of the p38 active form was relaxed with the energy minimization using the DISCOVER program [21]. The positions of the atoms in the loops and positions of all hydrogen atoms were first optimized, keeping the heavy atoms of the conserved regions as well as the atoms of $Mg^{2+}{}_2ATP^{-4}$ fixed until the root mean square (r.m.s.) gradient was smaller than 1 kcal/(mol Å). Next holding the conserved α carbon atoms fixed, positions of other atoms were optimized until the r.m.s. energy gradient was smaller than 0.1 kcal/(mol Å). Loops with the insertion and deletion regions were subjected to simulated annealing procedure from 2000 K to 300 K while keeping the remaining atoms of the model fixed, followed by additional energy minimization. A substrate septamer with the proposed following sequence, Arg-Glu-Asn-Thr-Glu-Tyr-Asp and residues located at positions: p – 3, p – 2, p – 1, p, p + 1, p + 2, p + 3, was modelled based on the PKI(5–24) inhibitor and structural properties of the p38 binding site. The side chains of the substrate were adjusted manually to maximize electrostatic and hydrophobic interactions with the p38 structure. The energy of the whole complex was minimized keeping the α-carbon atoms of the conserved regions, $Mg^{2+}{}_2ATP^{-4}$ and the substrate backbone heavy atoms fixed, or restrained by harmonic forces, until the r.m.s. energy gradient was 0.01 kcal/(mol Å). Finally all atoms were subjected to the energy minimization, except the restrained α carbon atoms of the conserved regions until the r.m.s. energy gradient was 0.001 kcal/(mol Å).

RESULTS AND DISCUSSION

The final model of the p38 active form containing $Mg^{2+}{}_2ATP^{-4}$ and the Arg-Glu-Asn-Thr-Glu-Tyr-Asp substrate is characterized by the r.m.s. deviation of the backbone atoms in the conserved regions equal to 2.22 Å as compared with the ERK2 active form model, and 2.77 Å as compared with the PKA crystal structure. Figure 3 presents a comparison of active forms of PKA and p38. The ATP-binding site in p38 is similar to the ATP-binding sites of PKA and ERK2. In the discussion below, the numbering in parentheses corresponds to the numbering of PKA, and without the parentheses it refers to a currently discussed kinase. There are two hydrogen bonds between the adenosine ring and the enzyme backbone atoms of p38 (see Fig. 4). The first is formed between N–H of Leu-108 (Val-123) and N_1 of adenosine, the second between C=O of Thr-106 (Glu-121) and N_6–H of adenosine. O_3–H of the ribose ring forms a hydrogen bond with C=O of Ser-154 (Glu-170). Phosphates are held by several groups. The interactions occur between NH_3^+ of Lys-53 (Lys-72) and α-phosphate, as well as between NH_3^+ of Lys-152 (Lys-168) and γ-phosphate. Mg^{2+} ions interact with COO^- of Asp-168 (Asp-184). Asn-155 (Asn-171) binds to the inhibitory Mg^{2+} ion and stabilizes the catalytic loop by a hydrogen bond to the α-carbonyl of Asp-150 (Asp-166). All hydrogen bonds between ATP and p38 are very similar to those seen in PKA or ERK2, i.e., they are formed by the same amino acids in the sequence alignment.

Figure 5 shows interactions between the proposed substrate molecule and its binding site on p38. Arg at the p – 3 site is located at the same position as in PKA. It submerges into enzyme and forms a hydrogen bond with the O3′ ribose oxygen atom of ATP. In ERK2,

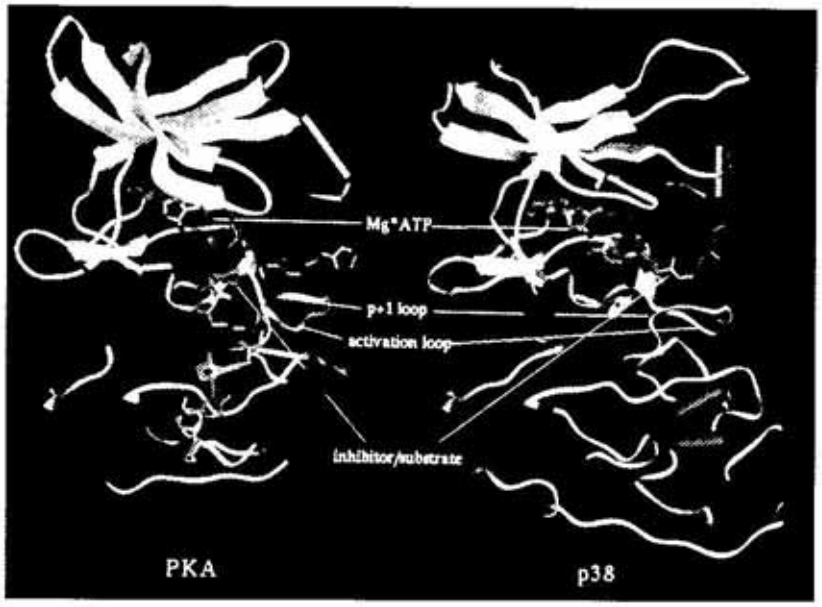

Figure 3. Ribbon diagrams of the PKA and p38 catalytic cores.

Both structures are complexed with ATP, as well as with an inhibitor in the case of PKA, and with the proposed substrate in the case of p38. The α-helices are in red, β-strands in yellow and connecting loops are in grey. The MgATP heavy atoms and the substrate backbone are represented by the stick model.

at the p – 3 site, a hydrophobic amino acid is preferred, and the O3′ atom interacts with Lys-112. At the p – 2 site a negative charged Glu is placed. It is located on the enzyme surface, and interacts with positive Lys-118. Glu at the p + 1 site interacts both with Arg--186 and Arg-189. Asp at p + 3 forms a pair with Lys-66. The Tyr ring at the p + 2 site is surrounded by several aromatic rings: Tyr-35, Phe-59 and His-64. The sequence p-(p + 1)-(p + 2) of the substrate, i.e., Thr-Glu-Tyr, is identical with that occuring in the activation loop of most MAP kinases. The latter may thus cause activation of the MAP kinases by themselves.

As mentioned above, there are nine conserved residues in the neighbourhood of active sites of the kinases, and in particular there are such residues in the p38 structure. On superimposing the upper domains of the

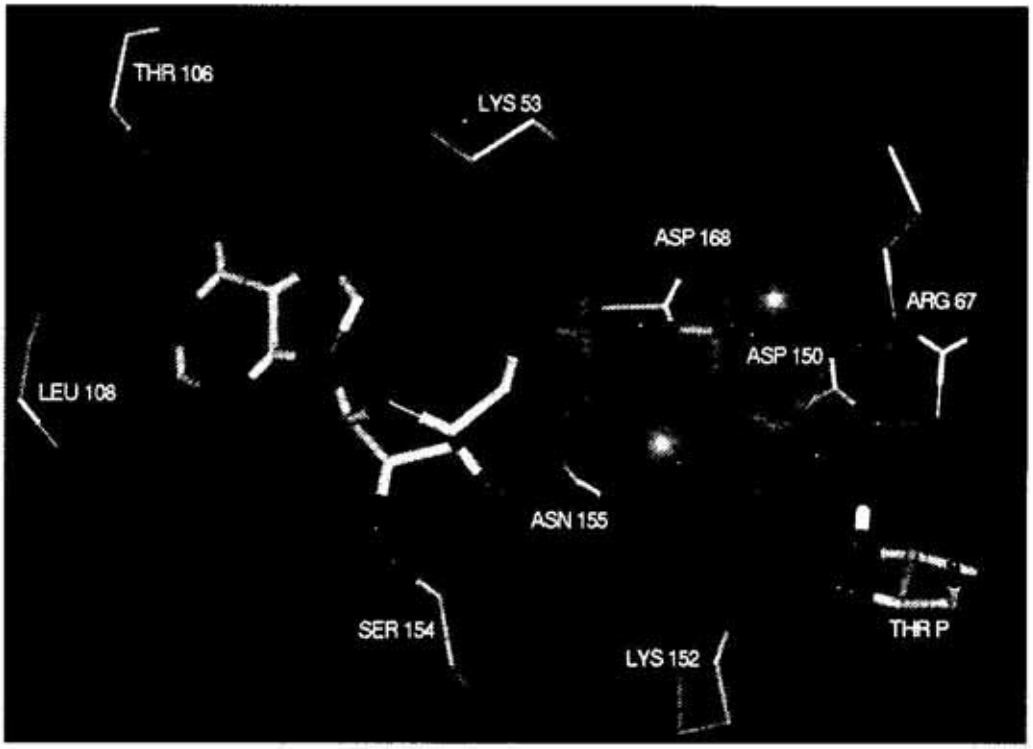

Figure 4. Stick model of p38 Mg-ATP binding site.

ATP and a fragment of the substrate molecule are drawn as bold sticks. Surrounding residues of the binding pocket are drawn as thin sticks. Mg^{2+} ions are represented by pink balls. Most important distances are given in Å.

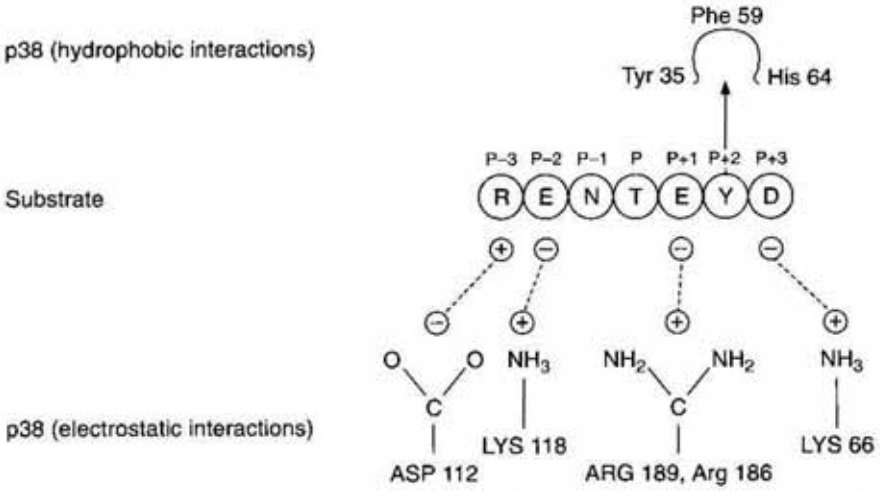

Figure 5. Contacts between p38 and its proposed substrate.

Charged residues of the substrate involved in its recognition by the kinase are shaded.

structures of p38 and PKA, one gets the r.m.s. deviation of 2.15 Å for the three conserved amino acids in that domain, superimposing of the lower domains leads to the r.m.s. of 1.10 Å for the other six residues. The values for p38 and ERK2 are equal to 1.25 Å and 0.87 Å, respectively. The r.m.s. values refer to the backbone atoms. The overlap of amino acids is not perfect, but their positions and hydrogen bonds, which they form with the substrate and ATP, are conserved. Since positions of the most important residues, such as Lys-152 (Lys-168) and Asp-150 (Asp-166) are very similar, the phosphorylation processes catalyzed by p38 and PKA are probably very similar. These conserved charged residues at the active site clearly play the same role in catalysis by p38 as in PKA [22]. Lys-152 (Lys-168) binds to the γ-phosphate helping to neutralize its charge and pulling the equilibrium in the direction of the closed conformation. The catalytic base Asp-150 (Asp-166), being 2.86 Å apart from the hydroxyl group of Thr, is positioning the hydroxyl for the nucleophilic attack on the γ-phosphate (cf. Fig. 4).

SUMMARY

Based on the crystallographically known active form of PKA and the inactive form of ERK2, the active form of p38 has been designed. The developed modelling procedure can also be applied to other kinases. The ATP-binding mode is very similar in the active forms of all three kinases studied. The active forms of p38 and PKA have very similar structures in the region where the substrate is bound. In particular, the separation distances between the ATP γ-phosphate and serine/threonine of the substrates are very similar both in p38 and PKA, which means that the transphosphorylation mechanism in these kinases should be also very similar. One should note, however, that p38 and ERK2 differ in the binding site region. Once the models of the active protein kinases are designed it will be possible to continue studying on several key problems, in particular on transphosphorylation mechanisms at a quantum level, as well as on design of inhibitors and potential drugs, competitive with ATP and/or substrates.

The computations were performed at ICM, Warsaw University, Poland. The authors thank Professor D. Shugar for stimulating discussions.

REFERENCES

1. Krebs, E.G. (1985) The phosphorylation of proteins: A major mechanism for biological regulation. *Biochem. Soc. Trans.* **13**, 813–820.
2. Knighton, D.R., Zheng, J., Ten Eyck, L.F., Ashford, V.A., Xuong, N.-H., Taylor, S.S. & Sowadski, J.M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407–414.
3. Knighton, D.R., Zheng, J., Ten Eyck, L.F., Xuong, N.-H., Taylor, S.S. & Sowadski, J.M. (1991) Structure of a peptide inhibitor bound

to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 414–420.

4. Zheng, J., Knighton, D.R., Xuong, N.-H., Taylor, S.S., Sowadski, J.M. & Ten Eyck, L.F. (1993) Crystal structures of the miristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations. *Protein Sci.* **2**, 1559–1573.

5. Zheng, J., Trafny, E.A., Knighton, D.R., Xuong, N.-H., Taylor, S.S., Ten Eyck, L.F. & Sowadski, J.M. (1993) 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and peptide inhibitor. *Acta Crystallogr. D* **49**, 362–365.

6. De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. & Kim, S.H. (1993) Crystal structure of cyclin-dependent kinase 2. *Nature* (London) **363**, 595–602.

7. Zhang, F., Strand, A., Robbins, D., Cobb, M.H. & Goldsmith, E.J. (1994) Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* (London) **367**, 704–711.

8. Su, S.-H., Parker, M.W., Lei, J.Y., Wilce, M.C.J., Benian, G.M. & Kemp, B.E. (1994) Insights into autoregulation from the crystal structure of twichin kinase. *Nature* (London) **369**, 578–581.

9. Hubbard, S.R., Wei, L., Ellis, L. & Hendrickson, W.A. (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* (London) **372**, 746–754.

10. Hanks, S.K., Quinn, A.M. & Hunter, T. (1988) The protein kinase family: Conserved feature and deduced phylogeny of the catalytic domains. *Science* **241**, 42–52.

11. Cox, S., Radzio-Andzelm, E. & Taylor, S.S. (1994) Domain movements in protein kinases. *Curr. Opin. Struct. Biol.* **4**, 893–901.

12. Bała, P., Grochowski, P., Lesyng, B. & McCammon, J.A. (1996) Quantum-classical molecular dynamics simulations of proton transfer processes in molecular complexes and in enzymes. *J. Phys. Chem.* **100**, 2535–2545.

13. Heidenreich, K.A. & Kumer, J.L. (1996) Inhibition of p38 mitogen-activated protein kinase by insulin in cultured fetal neurons. *J. Biol. Chem.* **271**, 9891–9894.

14. Raingeaud, J., Gupta, S., Rogers, J.S., Dickens, M., Han, J., Ulevitch, R.J. & Davis, R.J. (1995) Pro-inflammatory cytokines and environmental stress cause p38 mitogen-activated protein kinase activation by dual phosphorylation on tyrosine and threonine. *J. Biol. Chem.* **270**, 7420–7426.

15. Jiang, Y., Chen, C., Li, Z., Guo, W., Gegner, J.A., Lin, S. & Han, J. (1996) Characterization of structure and function of a new mitogen-activated protein kinase (p38β). *J. Biol. Chem.* **271**, 17920–17926.

16. Han, J., Lee, J.D., Bibbs, L. & Ulevitch, R.J. (1994) A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. *Science* **265**, 808–811.

17. Knighton, D.R., Pearson, R.B., Sowadski, J.M., Means, A.R., Ten Eyck, L.F., Taylor, S.S. & Kemp, B.E. (1992) Structural basis of the intrasteric regulation of myosin light chain kinases. *Science* **258**, 130–135.

18. Knighton, D.R., Cadena, D.L., Zheng, J., Ten Eyck, L.F., Taylor, S.S., Sowadski, J.M. & Gill, G.N. (1993) Structural features that specify tyrosine kinase activity deduced from homology modeling of the EGF receptor. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5001–5005.

19. *Homology User Guide, Version 95.0* (1995) Biosym Technologies, San Diego.

20. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr. & Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784.

21. *Discover User Guide, Version 95.0* (1995) Biosym Technologies, San Diego.

22. Madhusudan, E.A., Trafny, E.A., Xuong, N.-H., Adams, J.A., Ten Eyck, L.F., Taylor, S.S. & Sowadski, J.M. (1994) cAMP-dependent protein kinase: Crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci.* **3**, 176–187.