Acta Biochimica Polonica

Vol. 46 No. 3/1999

601-607

QUARTERLY

Modelling of insulin receptor tyrosine kinase in its active form: A case study for validation of theoretical methods°★

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

[1]*Department of Biophysics, Institute of Experimental Physics, Warsaw University, S. Żwirki i F. Wigury 93, 02-089 Warszawa, Poland;* [2]*Interdisciplinary Centre for Mathematical and Computational Modelling, Warsaw University, A. Pawińskiego 5a, 02-106 Warszawa, Poland*

Received: 21 July, 1999

Key words: insulin receptor kinase, protein tyrosine kinase, substrate specificity, phosphorylation, molecular modelling, molecular mechanics

An active form of an insulin receptor tyrosine kinase (IRK) catalytic core was modelled based on its experimentally known inactive form and the active form of a serine/threonine kinase, protein kinase A (PKA). This theoretical model was compared with the crystallographic structure of the active form of IRK reported later. The structures are very similar, which shows that all the most important features and interactions have been taken into account in the modelling procedure. The elaborated procedure can be applied to other tyrosine kinases. This would allow designing of a wide class of tyrosine kinase inhibitors, very important potential anti-cancer and/or anti-viral drugs.

Protein kinases comprise a large and diverse family of enzymes that play a key role in signal transduction [1]. Although they differ in size and subcellular distribution, all known protein kinases contain a two-domain conserved catalytic core. The "upper" and "lower" domains can form either "open" or "closed" enzyme conformation [2]. The closed conformation complexed with ATP and a peptide substrate is known to be the active form capable of transfering the γ-phosphate from ATP to the substrate. The full enzymatic activity requires also specific phosphorylation of the enzyme. Knowledge of structural details of the

★This study was supported by the Warsaw University within the projects BST-591/BF (K.G. and M.W.) and BST-641/99 (B.L.).

°Presented at the symposium "Conformation of peptides, proteins and nucleic acids" held in Rynia, Poland, on 26-29 May, 1999.

✉Corresponding author: e-mail: lesyng@icm.edu.pl; phone: (48 22) 874 9100; fax: (48 22) 874 9115.

Abbreviations: IRK, insulin receptor kinase; IR, insulin receptor; PDB, Protein Data Bank; PKA, protein kinase A; PKI, protein kinase inhibitor; AMP-PNP, 5'-adenylylimido-diphosphate; CDK2, cyclin-dependent kinase 2; r.m.s., root mean square; DSSP, Database of Secondary Structure in Proteins.

active forms of protein kinases is required for precise description of the enzymatic reactions as well as for rational design of enzyme's inhibitors. Modelling of the active form of insulin receptor tyrosine kinase (IRK) is the object of this study.

The insulin receptor (IR) is an $\alpha_2\beta_2$ transmembrane glycoprotein with an intrinsic protein kinase activity which, upon activation by insulin, mediates cell growth and metabolism [3–5]. Binding of insulin to the extracellular α chains of the receptor results in changes of the relative orientation of the β chains followed by *trans*-autophosphorylation (between two β chains) occurring in the cytoplasmic region of the protein [6]. Autophosphorylation of three tyrosine residues in the activation loop within the tyrosine kinase domain: Tyr-1158, Tyr-1162, Tyr-1163, is crucial for the kinase activity towards exogenous substrates [7, 8].

The IRK crystal structure, determined at a 2.1 Å resolution in its unphosphorylated inactive form, revealed a novel autoinhibition mechanism [9]. One of the tyrosines from the activation loop that is autophosphorylated in response to insulin, Tyr-1162, is bound in the active site, competing with the protein substrate. Residues at the beginning of the activation loop restrict the access to ATP, which prevents the *cis*-autophosphorylation of Tyr-1162. This intramolecular *trans*-autophosphorylation, upon the insulin binding, can occur only when Tyr-1162 is disengaged and MgATP is bound.

In order to understand at the molecular level the role of autophosphorylation for the kinase activity, to elucidate tyrosine kinase substrate specificity, and to gain an insight into the phosphotransfer mechanism, we constructed a 3-D model of the IRK catalytic core in its phosphorylated, active form complexed with MgATP and a peptide substrate, based on the previously elaborated procedure modelling active forms of protein kinases [10, 11]. When this modelling of IRK was completed, an experimental 1.9 Å crystal structure of IRK complexed with an ATP analogue and a peptide substrate has been reported [12]. This enabled comparison of the structures and provided a very useful test for validation of the applied theoretical methods and the modelling strategy.

METHODS

The theoretical model of the active ("closed"), phosphorylated form of the IRK catalytic core in complex with a peptide substrate and MgATP was designed, based on a 2.1 Å crystal structure of IRK in its inactive "open" form (PDB code: 1IRK) [9] and a 2.2 Å crystal structure of PKA in its active, "closed" form complexed with MnATP and a peptide inhibitor (PDB code: 1ATP) [13] (Fig. 1). The modelling procedure was carried out, according to the alignment shown in Fig. 2, using the INSIGHT II program package, version 95.0 [14].

An initial model of the IRK active form was prepared by changing the orientation of the upper lobe in the IRK inactive structure. Positions of the atoms in the lower lobe remained unchanged, and the upper lobe, treated as a rigid body, was transformed to obtain the same relative orientation of the lobes as observed in the structure of the PKA active form. The helix C was then rotated towards the C-terminal lobe to reconstruct a key interaction of Glu-1047 with Lys-1030. The linker between the domains was adjusted manually. The position of the activation loop (Gly-1149–Leu-1170) was remodelled based on plausible backbone fragments of the same length extracted from the PDB structures and the given IRK sequence. Three phosphate groups were attached to the side-chains of Tyr-1158, Tyr-1162 and Tyr-1163.

MgATP was then docked in the same orientation as MnATP in the PKA structure. Seven residues from the preferred IRK substrate sequence (EEDYMMM) [15] were modelled partially onto the backbone of the IRK pseudo-

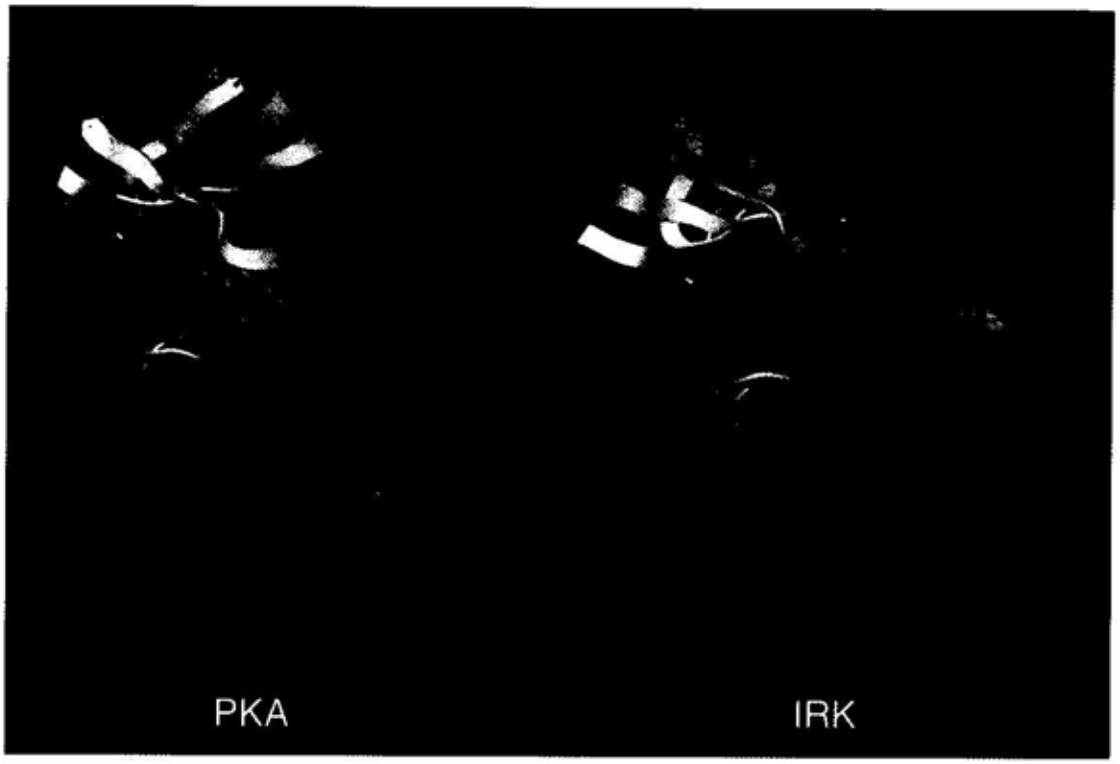

Figure 1. Ribbon diagrams of the protein kinase A (PKA) structure (PDB code: 1ATP, [13]) in its active form (complexed with MnATP and a protein kinase inhibitor) and the apo insulin receptor tyrosine kinase (IRK) structure (PDB code: 1IRK, [9] in its inactive form.

Secondary structure elements: α-helices and β-strands, assessed by DSSP [18], are shown in red and yellow, respectively. The α-helices are denoted by letters and the β-strands by numbers. MnATP is shown in gray and the protein kinase inhibitor as the blue ribbon.

substrate sequence (activation loop) taken from the structure of the IRK inactive form. Subsequently, the positions of several side-chains of the whole model were adjusted manually to remove unfavorable contacts and to maximize electrostatic and hydrophobic interactions. The model was then subjected to a series of energy minimization steps with the DISCOVER program [16] until the r.m.s. gradient was smaller than 0.001 kcal/(mol × Å). In addition, a simulated annealing procedure from 2000 K to 300 K was carried out to obtain an optimal structure of the activation loop region with the minimal energy. All energy optimizations were performed using the steepest descent and conjugate gradients methods using the AMBER force-field [17] with a distance dependent (4r) dielectric constant. The 1–4 nonbonded interactions were scaled by a factor of 0.5.

RESULTS

The model of the active form of the IRK catalytic core complexed with MgATP and a peptide substrate is the first theoretical protein tyrosine kinase structure in the active form. It allows for description of the reaction mechanism of this subfamily of protein kinases.

Our model shows that one of the most important residues responsible for the difference in the mechanism of catalysis between tyrosine and serine/threonine kinases is Arg-1136, which corresponds to Lys in serine/threonine kinases. However, Arg-1136 does not interact with the γ-phosphate of ATP, but stabilizes longer side-chain of tyrosine to be phosphorylated. The modelling studies show clearly that conformation of the P+1 loop, close to Pro-1172, conserved through the tyrosine

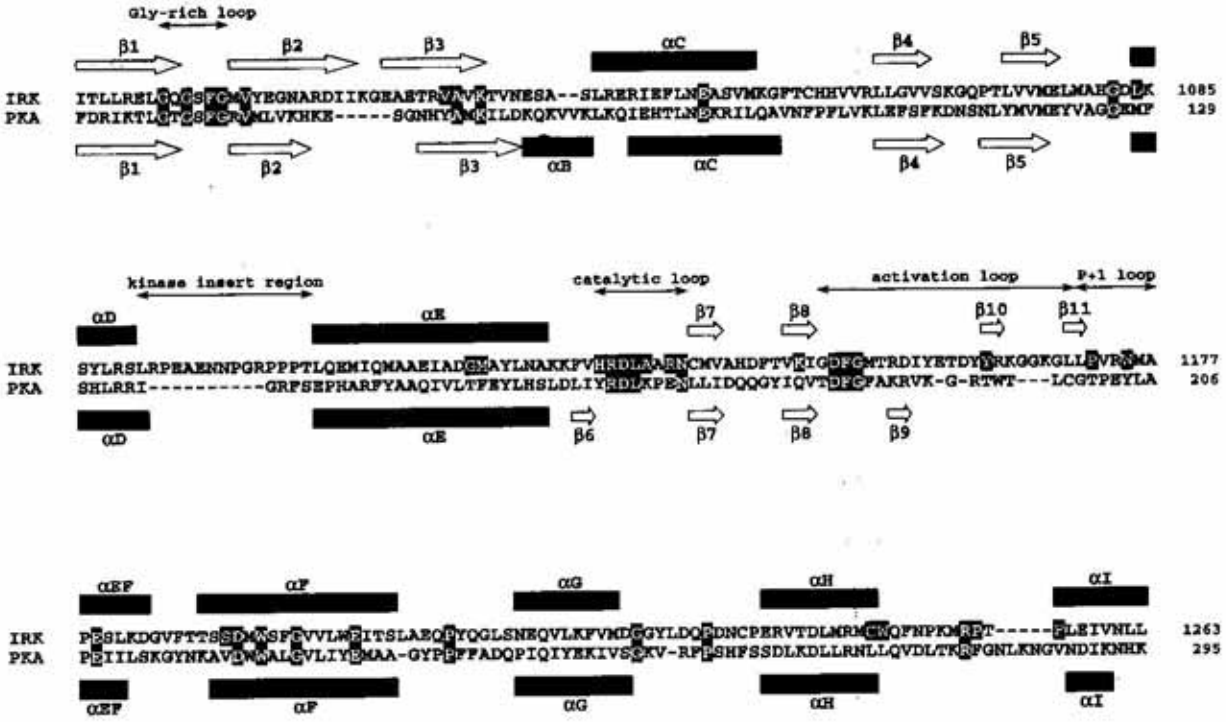

Figure 2. Structure-based sequence alignment of the catalytic core domains of IRK (PDB code: 1IR3, [12]) and PKA (PDB code: 1ATP, [13]).

Residues conserved within the tyrosine kinase subfamily are highlighted in blue, and those for all kinases are in green. Locations of the secondary structure elements in IRK and PKA assessed by DSSP [18] are marked above and beneath the sequences, respectively, β-strands are represented as yellow arrows, and α-helices as red rectangles. Positions of the phosphorylated residues: tyrosines in IRK and threonine in PKA, are marked by black dots.

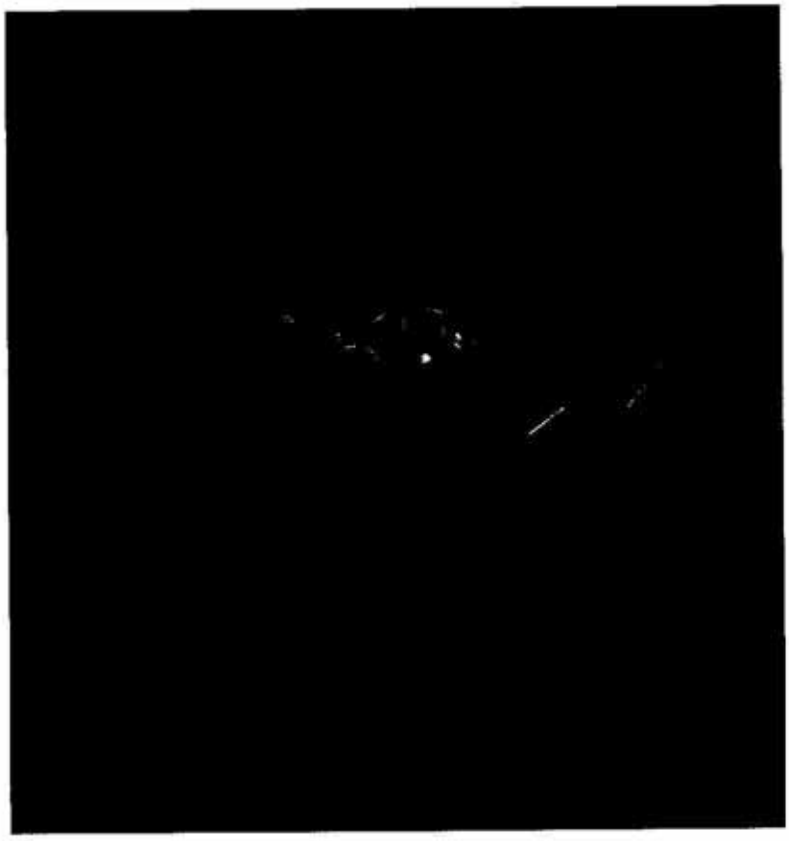

Figure 3. Comparison of the model and the crystal structure of IRK (PDB code: 1IR3, [12]) in its active conformation.

The catalytic cores of the model and the crystal structure are superimposed and represented as the α-carbon traces in red and blue, respectively. The MgATP/MgAMP-PNP and substrate α-carbon traces of the model and the crystal structure are shown in yellow and green, respectively.

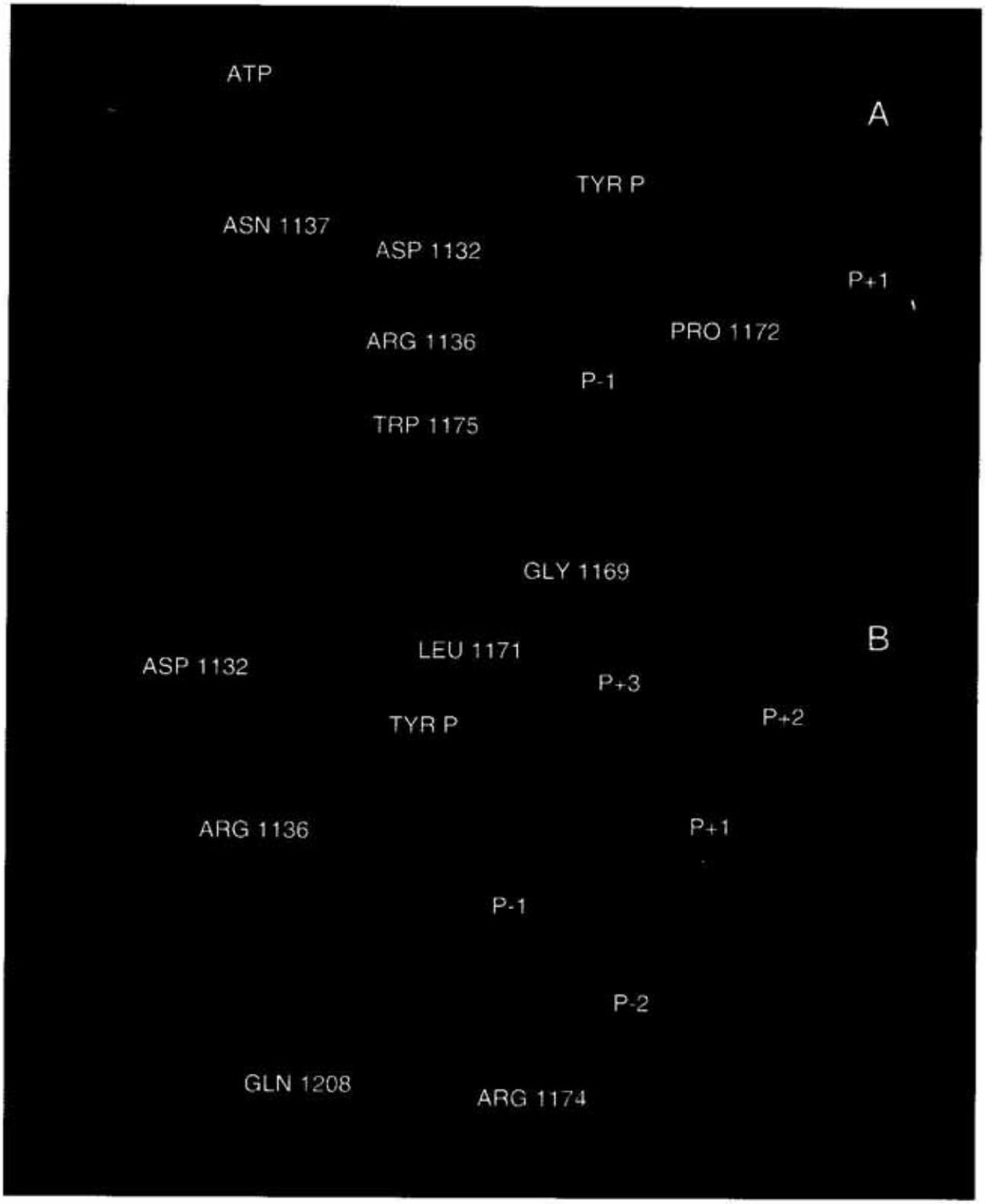

Figure 4. Comparison of the active site (A) and the substrate-binding region (B) for the model and the crystal structure (PDB code: 1IR3, [12]).

The most important residues responsible for the catalysis and the substrate-binding are represented by sticks and shown in orange and violet for the model and the crystal structure, respectively.

kinase subfamily, is the determinant of the tyrosine *versus* serine/threonine substrate selectivity. Furthermore, one clearly sees which residues contribute to the substrate-binding. This enables to understand the IRK substrate specificity and allows rational designing of tyrosine kinase inhibitors.

After these modelling studies have been completed, the experimental 1.9 Å crystal structure of the phosphorylated, activated form of IRK in complex with an ATP analogue and a peptide substrate has been reported [12]. This allowed comparison of the structures and enabled validation of the theoretical methods used in our modelling studies. A careful analysis shows that the modelled structure of the IRK active form was predicted correctly without any significant errors (Fig. 3). The r.m.s. deviation of 268 α-carbon positions in the catalytic core of the model with respect to the crystal structure is 2.3 Å. The lobe closure in the model is more tight than in the experimental structure. The ATP phosphate groups are slightly shifted in the direction of tyrosine to be phosphorylated. However, these differences may be due to the presence in the crystal structure of AMP-PNP, an ATP-analogue, precluding phosphorylation of the substrate.

The main differences are observed in the N-terminal part of the activation loop, which is very long and flexible. Some parts of this loop were taken from other kinases, but the position of the N-terminal part was generated *de novo*. The model predicts that Tyr-1163 plays the same role as Thr-197 in PKA, and its phosphorylation is required for the full activity of the kinase. One should note also that interactions of this phosphorylated Tyr-1163 with the other part of the enzyme are very similar in the model and crystal structures.

The structure of the active site and the substrate-binding region as well as positions of ATP and the peptide substrate have been predicted correctly (Fig. 4). Conformations of all key residues involved in the catalysis and the substrate-binding, especially Asp-1132, Arg-1136, Asn-1137, Trp-1175 and Tyr to be phosphorylated, correlate well with those observed in the experimental structure. Interactions of the peptide substrate with the residues on the IRK surface are also well reconstructed in the model.

SUMMARY

Our modelling studies together with the crystallographic structure of IRK [12] led to precise description of the IRK active form at the molecular level. The both structures of IRK in its active form provide an insight into the IRK receptor activation *via* autophosphorylation and explain the mechanism of phosphotransfer, the substrate specificity and tyrosine selectivity of this subfamily of protein kinases.

The detailed comparison of the model and the experimental crystal structure shows no essential discrepancies. Most of the amino-acid side-chains conformations, as well as the key interactions in the active site and the substrate-binding region, were predicted correctly. The main differences, observed in the activation loop, do not change the most important conclusions. The model structure reveals correctly that Tyr-1163, and not Tyr-1162, is structurally related to Thr-197 in PKA or Thr-160 in CDK2.

Our model is sufficiently accurate to allow formulation of reliable biological conclusions. In particular, this theoretical procedure can be applied to modelling of other protein tyrosine kinases as well as further rational design of their inhibitors as potential drugs.

REFERENCES

1. Hunter, T. (1991) Protein kinase classification. *Methods Enzymol.* **200**, 3–37.

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

3. White, M.F. & Kahn, C.R. (1994) The insulin signaling system. *J. Biol. Chem.* **269**, 1–4.

4. Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J.H., Masiarz, F., Kan, Y.W., Goldfine, I.D., Roth, R.A. & Rutter, W.J. (1985) The human insulin receptor cDNA: The structural basis for hormone-activated transmembrane signalling. *Cell* **40**, 747–758.

5. Ullrich, A., Bell, J.R., Chen, E.Y., Herrera, R., Petruzzelli, L.M., Dull, T.J., Gray, A., Coussens, L., Liao, Y.C., Tsubokawa, M., Mason, A., Seeburg, P.H., Grunfeld, C., Rosen, O.M. & Ramachandran, J. (1985) Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature* **313**, 756–761.

6. Wei, L., Hubbard, S.R., Hendrickson, W.A. & Ellis, L. (1995) Expression, characterization, and crystallization of the catalytic core of the human insulin receptor protein-tyrosine kinase domain. *J. Biol. Chem.* **270**, 8122–8130.

7. Rosen, O.M., Herrera, R., Olowe, Y., Petruzzelli, L.M. & Cobb, M.H. (1983) Phosphorylation activates the insulin receptor tyrosine protein kinase. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3237–3240.

8. Ellis, L., Clauser, E., Morgan, D.O., Edery, M., Roth, R.A. & Rutter, W.J. (1986) Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose. *Cell* **45**, 721–732.

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* **372**, 746–754.

10. Ginalski, K., Lesyng, B., Sowadski, J. & Wojciechowski, M. (1997) Modelling of active forms of protein kinases: p38 – A case study. *Acta Biochim. Polon.* **44**, 557–564.

11. Miller, M., Zbar, B., Ginalski, K. & Lesyng, B. (1998) Structural basis of oncogenesis caused by point mutations in the kinase domain of the MET proto-oncogene. *Cell. Mol. Biol. Lett.* **3**, 273–274.

12. Hubbard, S.R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.* **16**, 5572–5581.

13. 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 Crystallog. D* **49**, 362–365.

14. *INSIGHT II User Guide, Version 95.0* (1995) Biosym Technologies, San Diego.

15. Zhou, S., Carraway III, K.L., Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., Lorenzo, M.J., Ponder, B.A.J., Mayer, B.J. & Cantley, L.C. (1995) Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536– 539.

16. *DISCOVER User Guide, Version 95.0* (1995) Biosym Technologies, San Diego.

17. 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.

18. Kabsch, W. & Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.