Minireview

Protein kinases – structure and function

Dirk Bossemeyer*

Department of Pathochemistry, German Cancer Research Centre, INF 280, D-69120 Heidelberg, Germany

Received 10 May 1995

Abstract The solution of crystal structures from half a dozen protein kinases during the last four years in different laboratories has deepened our understanding of the catalysis and regulation of this enzyme class, and given a vigorous impetus to the whole field. Due to the great degree of sequence conservation among protein kinases the informational yield with every new structure is high, as each is a representative of the enzyme family in general and most often of a subclass in particular. This review will focus on the active site structure of cAMP-dependent protein kinase (cAPK) with special regard to two new crystal structures; one of an active protein kinase CK1*, which may represent an as yet unsolved step in the kinetic pathway, and the other of the insulin receptor kinase domain, the first structure of a tyrosine kinase.

Key words: Crystal structure; cAMP-dependent protein kinase; Protein kinase CK1; Insulin receptor; Catalytic site; Conserved sequence motif

1. Introduction

Protein kinases are critically involved in almost every regulatable cellular process. Several hundreds of members of the protein kinase family all possess a sequence of eleven conserved subdomains, the catalytic core, in conjunction with their shared task of modulating substrate proteins by phosphorylating serine (S), threonine (T) or tyrosine (Y) residues [1]. Until very recently all protein kinase crystal structures have been from members of the ST-subfamily; for reviews see [2-4]. Now Hubbard et al. have filled this gap, and reported a structure of the protein tyrosine kinase (PTK) domain of the insulin receptor IRK [5]. The various crystal structures confirm the structural conservation of the protein kinase catalytic core, at the same time demonstrating how protein kinases reconcile the high degree of conservation essential for the catalytic function with the need for individual solutions for their diverse biological roles. As a rule, conservation in sequence indicates conserved spatial structure and function, while individual properties relate to residues, loops or insertions not conserved in the family as a whole. However, exceptions are apparent. In IRK one of the most highly conserved sequence motifs, the triad Asp-Phe-Gly in subdomain VII, is displaced and part of an autoinhibitory mechanism. On the other hand, protein kinase CK1, the structure of which has been solved recently [6], lacks the conserved motif APE in subdomain VIII and a conserved arginine residue in subdomain XI without significant conformational restrictions. More surprises may be ahead.

2. The catalytic situation in PKs

The most comprehensive set of structural data on a single enzyme has been accumulated for porcine and recombinant mouse catalytic subunit of cAPK, which provided the first crystal structure of a protein kinase [7], and still is the only protein kinase which has been crystallised in ternary complexes, presumably showing the catalytic conformation with nucleotide and (pseudo)substrate bound [8-10]. Yet, we have good reason to assume that during catalysis all protein kinases will adopt a conformation very similar to that of the ternary complex of cAPK, making the cAMP-dependent PK a universal structural reference.

One or several crystal structures were also reported from cyclin dependent kinase CDK2 [11], the extracellular signal regulated kinase ERK2 [12,13] and the myosine light chain kinase twitchin [14]. The inactive states of these enzymes revealed many important features of enzyme regulation and specificity [2-4]. The recently reported structures of CK1l and IRK now are from potentially active enzymes, in addition to providing new insights into the specific and unique features of each enzyme they allow to evaluate the active site in comparison to that of cAPK more closely.

2.1. The catalytic protein kinase core

Protein kinases are highly flexible enzymes, large rotational movements of the conserved kinase lobes [15] as well as movements of flexible loops and domains accompany binding of substrates, cofactors, auto inhibitory domains, or interacting proteins, during catalysis or for regulation. The enzyme is bifold with the active site located in the mouth region of a deep cleft between the lobes (Fig. 1). This cleft completely accommodates the ATP molecule, with the γ-phosphate oriented outwards. The protein substrate binding site is vis-à-vis at the opening of the cleft. Most of the highly conserved residues cluster in this active site, approaching from different parts of both lobes.

2.2. The N-lobe

The smaller amino terminal N-lobe is dominated by a five stranded antiparallel β-sheet (β1-β5) and a long α-helix (C). In cAPK an additional helix B inserts between strand β3 and helix C. In some PKs, including cAPK, this region is used for pro-

*Corresponding author. e-mail bossemeyer@dkfz-heidelberg.de

Abbreviations: cAPK, cAMP-dependent protein kinase; CK1, protein kinase CK1, formerly casein kinase I; CK1l, variant of CK1; ERK, extracellular signal regulated kinase; IRK, kinase domain fragment of insulin receptor; CDK2, cyclin dependent kinase 2. Residue numbers of homologous residues of cAPK are in brackets.
tein-protein interactions. In cAPK His87 at the N-terminus of helix C contacts the C-terminal lobe via the constitutive phosphorylated Thr197. This contact is lost in the open conformation of the enzyme [15,16]. Interaction of His87 with the regulatory R1-subunit is indicated by complementary effects of mutations in His87 of the C-subunit and Ser99 in the R1-subunit [17]. Cyclin binding, a precondition for the activation of CDKs, is affected by mutations of residues in the region...
between 3 and helix C in CDK [18]. The first two \(\beta\)-strands contain the glycine-rich sequence GXGXXGXXV, a turn connecting both strands is formed by the central four amino acids of this motif. The conserved glycine-rich sequence contributes in many ways to protein kinase function [19]. The invariant conserved glycines allow a close approach of the nucleotide to the peptide backbone. From the third \(\beta\)-strand an invariant lysine, Lys72 in cAPK, contacts \(\alpha\) and \(\beta\)-phosphoryl groups of the bound ATP. Protein kinases are most sensitive to mutations of this lysine, which probably contributes to the correct stereochemical orientation of the triphosphate and the \(\gamma\)-phosphate for catalysis [20,21]. The spatial position of Lys72, and in turn that of the bound phosphoryl groups, is secured by a salt bridge to Glu91, a residue in the middle of helix C. It is assumed that the interaction between Lys72 and Glu91 is conserved in the catalytic conformation of PKs. This salt bridge is missing in some structures of inactive kinases, and also in the potentially active IRK: in CDK2, helix C is rotated, and Glu51 points to the solvent. Lys33 (Lys72 in cAPK) interacts with the \(\alpha\)-phosphate only, the \(\beta\)-phosphate is oriented differently. Cyclin binding might allow the active site residues and the triphosphate to achieve the active conformation found in cAPK. In the twitchin kinase, Lys5971 (Lys72) interacts with residues of a C-terminal intrasteric autoinhibitory peptide, which traverses the active site and is assumed to dislodge on calmodulin binding [22].

2.3. The C-lobe

Several residues of the larger helical C-terminal lobe interact with the triphosphate group of ATP. The conserved core of this lobe has seven \(\alpha\)-helices in cAPK, and four short \(\beta\)-strands, with a central bundle of four antiparallel \(\alpha\)-helices. The \(\beta\)-strands \(\beta6\) to \(\beta9\) form the bottom of the cleft beneath the adenosine moiety. The catalytic loop between \(\beta6\) and \(\beta7\), and the metal binding loop between \(\beta8\) and \(\beta9\) are of primary importance for the active site of PKs. The catalytic loop from Arg165 to Asn171 carries residues which are directly involved in catalysis. Here resides the catalytic base, Asp166, which is believed to abstract the proton from the hydroxyl group of the substrate amino acid. Of special significance is Lys168. Its side-chain contacts the \(\gamma\)-phosphate, neutralises the negative charge, and stabilises the presumed intermediate state. This residue is conserved in kinases with ST-specificity, including members of the STY-kinases [23]. In PTKs an arginine in equivalent position or two residues further in the primary sequence (Glu170 in cAPK) has been predicted to take up the functional role of this lysine [9].

The metal-binding loop contains the conserved motif DFG (Asp184PheGly). In cAPK the carbonylate of Asp184 serves as a ligand of the metal ion M2 which co-ordinates \(\beta\) and \(\gamma\)-phosphoryl oxygens, and a pair of water molecules. Two metal ions are found liganded to ATP; one is essential for activity, the other is inhibitory and binds with low affinity. Because of its nearly optimal octahedral co-ordination scheme, and the relative number of charged ligands, the M2 site is assumed to be the primary metal binding site. Asp184 may serve to establish the stereochemistry of the triphosphate group and to neutralise charge at the \(\gamma\)-phosphate via the metal ion. Mutating this residue results in inactive kinase [24]. The secondary metal ion co-ordinates the \(\alpha\)- and \(\gamma\)-phosphoryl groups, a water molecule, and the sidechain of invariant Asn171 at the end of the catalytic loop. The co-ordination scheme of this metal ion is trigonal bipyramidal, if one of the oxygens of Asp184 at a distance of 2.6 \(\AA\) (compared to about 2.0 \(\AA\) for the metal co-ordination sphere) is taken as a further weak ligand. Both metal ions potentially contribute to charge dispersal at the \(\gamma\)-phosphate, which is supported by kinetic data which relate the inhibitory effect of the secondary metal ion largely to a reduction in release of ADP, with little effect on the actual phosphoryl transfer reaction [25,26].

3. IRK

IRK is the fragment of the cytoplasmic kinase domain of the insulin receptor \(\beta\)-chain, an \(\alpha_2\beta_2\) heterotetramer, which is assumed to undergo a change in quaternary structure on insulin binding to the \(\alpha\)-chain, leading to activation of the kinase and trans-autophosphorylation of a number of tyrosine residues in the \(\beta\)-chain [5,27]. The overall structure shows great similarity to that of the known structures of ST-kinases, confirming the universal validity of the catalytic core. IRK, although unphosphorylated, is potentially active, and becomes auto-phosphorylated upon addition of MgATP. The enzyme is in an open conformation, which is stabilised by steric interactions between the glycine rich sequence and the DFG motif (Fig. 1). The open conformation leads to relative displacement of catalytic residues from the small lobe compared to those of the large lobe, and also does not allow formation of the salt bridge between Lys1030 (Lys72) and Glu1047 (Glu91). Clearly rigid body rotations and in addition conformational changes are necessary for IRK to take up the catalytic conformation upon activation.

One of the most interesting aspects of this tyrosine kinase structure is the mechanism of autoinhibition it revealed. Many kinases use the loop between DFG (subdomain VII) and APE (subdomain VIII) for regulation, it differs, however, largely in size and conformation in each enzyme [2]. In IRK this loop contains the three autophosphorylation site tyrosines. We have the intriguing situation of finding one of them, Tyr1162, in the active site, occupying a position seemingly poised for auto-phosphorylation. However, cis-autophosphorylation cannot occur because the ATP binding site is efficiently blocked by Phe1151 of the DFG motif (Fig. 1). Apparently, "self" binding of the activation loop transmits serious distortions to the DFG motif. Hubbard et al. suggest trans-autophosphorylation dislodges the autoinhibitory loop, in agreement with biochemical findings [28,29].

Although not in the catalytic conformation, the structure of IRK allows examination of interactions of the residues of the catalytic loop. The hydroxyl group of Tyr1162 is hydrogen bonded to the carboxylate of Asp1132 (Asp166), supporting the role of this residue as the catalytic base in the kinase reaction [8,9,30], in contradiction to the results of others [31]. In addition, the hydroxyl group has contact to the guanidinium group of Arg1136 (168 + 2), as a direct evidence for the proposed functional replacement of the catalytic loop lysine Lys168 (cAPK) by arginine. Superimposition of the kinase structures shows similar locations of the guanidinium group of this arginine and the ammonium group of Lys168 in cAPK (Fig. 2). Apparently a positive charge is conserved throughout the protein kinase family in three-dimensional space, rather then in the primary sequence.
The bound autophosphorylation substrate also provides an explanation how tyrosine selectivity is determined. The rather long sidechain of a tyrosine residue compared to serine or threonine is compensated for by a different position of its main chain, governed by PTK specific residues of the P + 1 loop.

The IRK mechanism of autoinhibition by a real substrate, instead of a pseudosubstrate, is novel, and may apply for other PTKs as well [5]. However, inhibition by a bona fide substrate is also known from cAPK. The R11-subunit inhibits the C-subunit in the absence of ATP, but becomes readily phosphorylated in the presence of ATP, with little change in the inhibitory potential [32], adding an 'inhibitory product' to the varied collection of protein kinase regulatory mechanisms.

4. CKII

The crystal structure of a truncated variant, CKII, of protein kinase CK1, a kinase ubiquitous to all eukaryotes, in its MgATP complex has been solved [6]. The protein is unphosphorylated, though active and the peptide binding site is fully accessible. Two sulfate ions from the crystallisation medium show potential phosphate binding sites, one of them being homologous to the Thr197 phosphate binding site in cAPK. The CKII substrate selectivity is directed towards phosphate groups N-terminal of the P-site. The other sulfate ion, bound in the cleft, seems to indicate the binding site for this substrate recognition phosphate. The overall architecture resembles the known protein kinases, with largest differences observed in surface loops. Again it is a loop between subdomain VII and VIII (L-9D), which contains three potential phosphate oxygens as via the metal ion in order to stabilise the triphosphate conformation.

The solution of protein kinase crystal structures in the future may complete not only our present picture of the enzymatic pathway, but very likely will give us also new insights into more possible variations on a common conserved catalytic core.

Acknowledgements: Atomic coordinates of CKII were kindly provided by Xiaodong Cheng, the coordinates of IRK were taken from Brookhaven Protein Data Bank entry pdblirk.ent. I should like to express my thanks to Volker Kinzel and Robert Huber for continuous support.

References