Article No. jmbi.1999.2973 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) **291**, 239–247

JMB

COMMUNICATION

Phosphatidylinositol Phosphate Kinase: A Link between Protein Kinase and Glutathione Synthase Folds

Nick V. Grishin*

National Center for Biotechnology Information
National Library of Medicine
National Institutes of Health
Bldg. 38A, Bethesda, MD
20894, USA

Comparisons of serine/threonine protein kinase (PK) and type IIβ phosphatidylinositol phosphate kinase (PIPK) structures with each other and also with other proteins reveal structural and functional similarity between the two kinases and proteins of the glutathione synthase fold (ATP-grasp). This suggests that these enzymes are evolutionarily related. The structure of PIPK, which clearly resembles both PK and ATP-grasp, provides a link between the two proteins and establishes that the C-terminal domains of PK, PIPK and ATP-grasp share the same fold. The functional implications of the proposed homology are discussed.

Keywords: protein kinase; lipid kinase; ATP-grasp fold; fold evolution; SAICAR synthase

The structure of type IIβ phosphatidylinositol phosphate (PIP) kinase (PIPK) revealed a similarity to protein serine/threonine and tyrosine kinases (PKs), that is not easily recognizable from their sequences (Rao *et al.*, 1998). Both PKs, for example, cAMP-dependent protein kinase (CAPK, 1cdk†; Knighton *et al.*, 1991a,b; Bossemeyer *et al.*, 1993) and PIPKs (1bo1; Rao *et al.*, 1998) are two-domain α+β proteins with an ATP-binding site embedded within the interdomain cleft. The structural similarity between the two kinase classes resides mainly in the cleft region and includes the arrangement of the ATP-binding residues (Rao *et al.*, 1998).

The similarity, however, is not localized to a few patches of structural elements, since the topology of the N-terminal domain core, which consists of a five-stranded β-sheet and an α-helix, is identical for both PKs and PIPK (Figure 1(a) and (b)). The DALI algorithm (Holm & Sander, 1997) aligns C^{α} atoms of 120-150 residues of PIPK with different PKs with root-mean-square deviations (RMSD) of 3.6-4.8 Å and *Z*-scores of 3.3-5.4. The major differences between the two kinase structures are confined to the C-terminal domain, which is mainly α-helical in PKs (Figure 1(a)) but is largely β-structural in PIPK (Figure 1(b)).

Nevertheless, in addition to the common ATP-binding site, the C-terminal domains of kinases from both classes share two unusual structural features: a left-handed ααββ unit (DEij on Figure 1(a) and (b)) and crossing loops (a loop between E and i and a loop between j and F on Figure 1(a) and (b)). The left-handed character of DEij leads to the strand f being hydrogen-bonded to the strand i. In other structures, additional β-strands usually have been found to extend the β-sheet in a left-handed αββ unit (Kajava, 1992) (Eij, shown in red in Figure 1) beside the β-strand j rather than beside β-strand i, as observed for PKs and PIPK. Due to the unusual crossover connection, α-helix F is positioned between α-helices D and E. These structural and functional similarities suggest that protein and PIP kinases are homologous (Rao *et al.*, 1998) rather than the results of structural convergence.

As for the differences between the two kinase superfamilies, the β-strands are longer in the

---

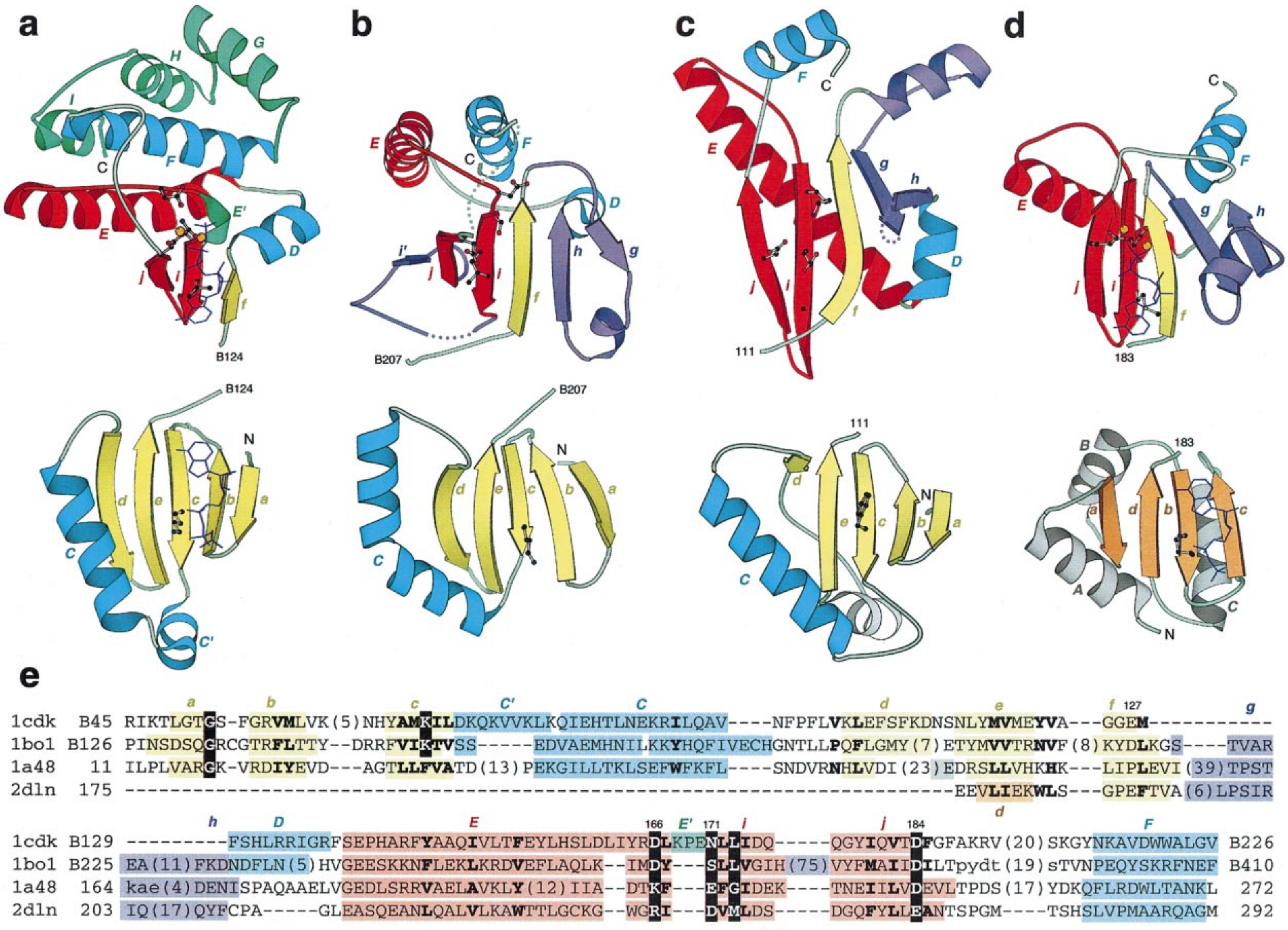
Abbreviations used: PK, serine/threonine or tyrosine protein kinase; cAPK, cAMP-dependent protein kinase; AK, aminoglycoside kinase; PIPK, type IIβ phosphatidylinositol phosphate kinase; DD-ligase, D-ala-D-ala ligase; ANP, adenylyl imidodiphosphate; SAICAR, phosphoribosylaminoimidazolesuccinocarboxamide; PIP, phosphatidylinositol phosphate; RMSD, root-mean-square deviation; PDB, Protein Data Bank; PHY, 1(*s*)-aminoethyl-(2-carboxy-1-propyl)-phosphoryl-phosphoric acid.

---

† Protein Data Bank (Abola *et al.*, 1997) (PDB) ID codes are given when available.

E-mail address of the corresponding author: grishin@ncbi.nlm.nih.gov

0022-2836/99/320239–9 $30.00/0



```
1cdk  B45 RIKTLGTGS-FGRVMLVK(5)NHYAMKILDKQKVVKLKQIEHTLNEKRILQAV----NFPFLVKLEFSFKDNSNLYMVMEYVA---GGEM-----------
1bo1 B126 PINSDSQGRCGTRFLTTY--DRRFVIKTVSS-----EDVAEMHNILKKYHQFIVECHGNTLLPQFLGMY(7)ETYMVVTRNVF(8)KYDLKGS----TVAR
1a48   11 ILPLVARGK-VRDIYEVD---AGTLLFVATD(13)PEKGILLTKLSEFWFKFL----SNDVRNHLVDI(23)EDRSLLVHKHK---LIPLEVI(39)TPST
2dln  175 -------------------------------------------------------------------------EEVLIEKWLS---GPEFTVA(6)LPSIR

1cdk B129 ----------FSHLRRIGRFSEPHARFYAAQIVLTFEYLHSLDLIYRDLKPENLLIDQ-----QGYIQVTDFGFAKRV(20)SKGYNKAVDWWALGV B226
1bo1 B225 EA(11)FKDNDFLN(5)HVGEESKKNFLEKLKRDVEFLAQLK--IMDY---SLLVGIH(75)VYFMAIIDILTpydt(19)sTVNPEQYSKRFNEF B410
1a48  164 kae(4)DENISPAQAAELVGEDLSRRVAELAVKLY(12)IIA--DTKF---EFGIDEK----TNEIILVDEVLTPDS(17)YDKQFLRDWLTANKL  272
2dln  203 IQ(17)QYFCPA-----GLEASQEANLQALVLKAWTTLGCKG--WGRI---DVMLDS-----DGQFYLLEANTSPGM----TSHSLVPMAARQAGM  292
```

C-terminal domain of PIPK than in PKs, and they comprise most of the structure (Figure 1(b), top), including additional β-strands g, h and i′, that are absent in PKs. Indeed, the β-sheet in the C-terminal domain of PKs is barely recognizable (Figure 1(a), top) resulting in annotation of the C-terminal domain of PKs ''all alpha'' in SCOP (Hubbard *et al.*, 1997) and ''mainly alpha'' in CATH (Orengo *et al.*, 1997). The PIPK structure emphasizes the importance of the β-sheet in the C-terminal domain of the kinase fold. Comparison of PK and PIPK structures reveals the core elements of the C-terminal domain of the kinase fold, namely the three-stranded antiparallel β-sheet (f, i and j), whose hydrophobic core is completed with the helices E and F. An ATP molecule binds along the β-sheet, forming most of its interactions with the central strand i.

All PKs possess highly similar structures. Even the most divergent member of the PK fold, namely aminoglycoside kinase (AK) (Hon *et al.*, 1997) shares extensive structural and functional similarity with the PKs. An RMSD value of 1.8 Å for a comparison of core regions of AK and cAPK is comparable to the values among different PKs (Hon *et al.*, 1997). Recently, AK has also been shown to have protein kinase activity (Daigle *et al.*, 1999). Homology between AK and PK can also be established from their sequences using the recently developed profile search tool PSI-BLAST (Altschul *et al.*, 1997). Using the AK sequence (SwissProt ID: KKA3_ENTFA) as a query for PSI-BLAST with default parameters, the first sequence of a PK (namely RKN2_MYXXA) is identified by the second iteration with a score of 45 bits and an E-value of 8×10^{-4}. In contrast, PIPK was not detectable in these searches.

PKs are rather isolated from all other folds. Addition of a divergent member, such as PIPK, to the PK fold might facilitate detection of evolutionary links with other proteins. Since the C-terminal domain of the kinase fold demonstrates some unusual features as discussed above, the proteins that share these features are likely to be homologs (Murzin, 1998). When the number of secondary structural elements in a protein is small, it is more likely that examples of given topology arose independently in evolution. However, different spatial arrangements are not equally probable due, in part, to the chiral properties of amino acids. In general, significance of the match rises inversely to the number of different protein families that share the particular supersecondary structural element.

A set of superfamilies was selected from SCOP (Hubbard *et al.*, 1997) according to the following criteria: the presence of (i) a left-handed αββ unit (Kajava, 1992) (with elements designated Eij) and (ii) a β-strand (designated f) located outside the sequence boundaries set by Eij, and hydrogen-bonded to the β-strand i in an antiparallel fashion (as in Figure 1(a) and (b)). Surprisingly, this particular packing of the four secondary structural elements occurs in only nine protein superfamilies out of about 600 superfamilies. Of these, two are related to the remaining seven by circular permutation. Among the chosen representatives of the nine superfamilies, only the C-terminal domain of the glutathione synthase fold (Hara *et al.*, 1996) (1gsa) contains helix F and binds ATP along the β-sheet formed by the strands f, i, and j. The glutathione synthase fold, called ATP-grasp in SCOP (Hubbard *et al.*, 1997), comprises several enzyme families, including D-ala-D-ala ligase (Fan *et al.*,

Figure 1. Comparison of kinase and glutathione synthase folds. ''Open book'' ribbon diagrams of PK, PIPK, SAICAR synthase, and D-ala-D-ala ligase structures are drawn using the program MOLSCRIPT (Kraulis, 1991). The N-terminal domain (bottom row) and the C-terminal domain (top row) are separated from each other for the purposes of clarity. The domain interface (ATP-binding cleft) is facing up (open book). The N and C termini of the displayed segment of a protein chain are labeled in roman upper case black letters as well as the numbers of the last and the first residue in the N-terminal and the C-terminal domain, respectively. Secondary structural elements are labeled by color-coded italicized letters: lower case is for β-strands and upper case is for α-helices with letter colors corresponding to the colors of the structural element. The αββ unit, essential for ATP-binding is shown in red. ''Cofactor'' is shown in a dark blue wire presentation and is indicated on each domain to mark the binding site. Metal cations are shown as orange balls. Catalytic, ATP and metal-binding residues are shown in a ball-and-stick representation. (a) cAMP-dependent protein kinase (Bossemeyer *et al.*, 1993) in complex with adenylyl imidodiphosphate and Mn^{2+} (1cdk, residues B46-B298). The additional helices are shown in green. (b) Type IIβ phosphatidylinositol phosphate kinase (Rao *et al.*, 1998) (1bo1, residues B126-B416). (c) SAICAR synthase (Levdikov *et al.*, 1998) (1a48, residues 11-272). (d) D-ala-D-ala ligase (Fan *et al.*, 1994) in complex with ADP and Mg^{2+} (2dln, residues 95-292). The N-terminal domain of ATP-grasp fold (the ''central'' domain in the protein) is colored differently to emphasize the fold difference. The additional structural elements in the C-terminal domain (compared to PKs) are shown in purple in (b), (c) and (d). Dots stand for the disordered regions. (e) Structure-based sequence alignment of 1cdk, 1bo1, 1a48, and 2dln. The alignment was constructed using DALI (Holm & Sander, 1997) and corrected manually to incorporate additional elements and adjust several regions. Color and letter-coding of structural elements in the sequence is as described for (a), (b), (c) and (d). The chain ID (where available) and the residue number of the first amino acid in each line and the last amino acid in the alignment are indicated. Numbers in parentheses interrupting the sequence indicate the number of residues in insertions not displayed in the alignment. Residues disordered in the structure are shown in lower case letters. Residues that form the conserved hydrophobic core are shown in bold letters. Residues that participate in catalysis, ATP and metal cation binding (displayed in ball-and-stick models in (a), (b), (c) and (d)) and catalysis are shown in white on black. cAPK residue numbers mentioned in Table 1 appear above the alignment.

1994) (DD-ligase, 2dln; Figure 1(d)). DD-Ligase was chosen as an ATP-grasp fold representative, since its ATP-binding site has been compared to cAPK (Denessiouk *et al.*, 1998; Kobayashi & Go, 1997)).

Other proteins (Figure 2) which contain the four core elements with an arrangement similar to the kinase C-terminal domain (β-strands f, i, j and α-helix E on Figure 1) are D-amino acid aminotransferase N-terminal domain (Sugio *et al.*, 1995) (Figure 2(a)), cystatins (Bode *et al.*, 1988) (Figure 2(b)), two DNA-binding proteins, namely replication terminator protein (Tus) C-terminal domain (Kamada *et al.*, 1996) (Figure 2(c)) and cro λ repressor (Albright & Matthews, 1998) (Figure 2(d)), MHC antigen-recognition domain (Vaghn & Bjorkman, 1998) (Figure 2(e)) and N-terminal fragment of Vaccinia DNA topoisomerase I (Sharma *et al.*, 1994) (Figure 2(f)). With circular permutation, the same elements are present in some actin-depolymerizing proteins, for example, villin (Markus *et al.*, 1997) (Figure 2(g)) and in the C-terminal domain of Fe-Mn superoxide dismutase (Ludwig *et al.*, 1991) (Figure 2(h)). None of these protein domains binds ATP and none utilizes the ij hairpin for binding any other cofactor in similar manner to protein/lipid kinases and ATP-grasp.

Thus, of approximately 600 superfamilies only nine passed the structure test and of these only one superfamily passed the function test as representing a possible homolog of PKs. The structure-function comparison of PK and glutathione synthase folds was undertaken subsequently in order to investigate the homology prediction. Structural similarities in the ATP-binding regions were studied previously for cAPK and DD-ligase (Denessiouk *et al.*, 1998; Kobayashi & Go, 1997). Indeed, the active site that includes several ATP and Mg^{2+}/Mn^{2+}-binding residues displays strong resemblance between ATP-grasp and both PK and PIPK superfamilies (Figure 1). Superposition of the active site regions from the C-terminal domains of cAPK, PIPK and DD-ligase is shown in Figure 3. The structures were superimposed to minimize RMSD values between backbone atoms of selected segments (Figure 3). Unexpectedly, superposition revealed that the active site region of cAPK is significantly more similar to DD-ligase than it is to PIPK (Figure 3). Superposition of 84 backbone atoms resulted in an RMSD value of 1.07 Å between cAPK and DD-ligase compared with RMSD values of 2.95 Å and 2.97 Å for PIPK - cAPK and PIPK - DD-ligase comparisons, respectively. The same trend is observed for the distances between functionally equivalent side-chain atoms (Table 1A). The result is surprising, since although there is little doubt that cAPK and PIPK are homologs, their structures display larger RMSD values. In contrast, cAPK and DD-ligase, whose homology relationship is less evident, possess nearly identical backbone structure and highly similar side-chain orientations. Consequently, although DD-ligase is more similar to PIPK globally due to its large β-sheet, DD-ligase is more similar to cAPK locally

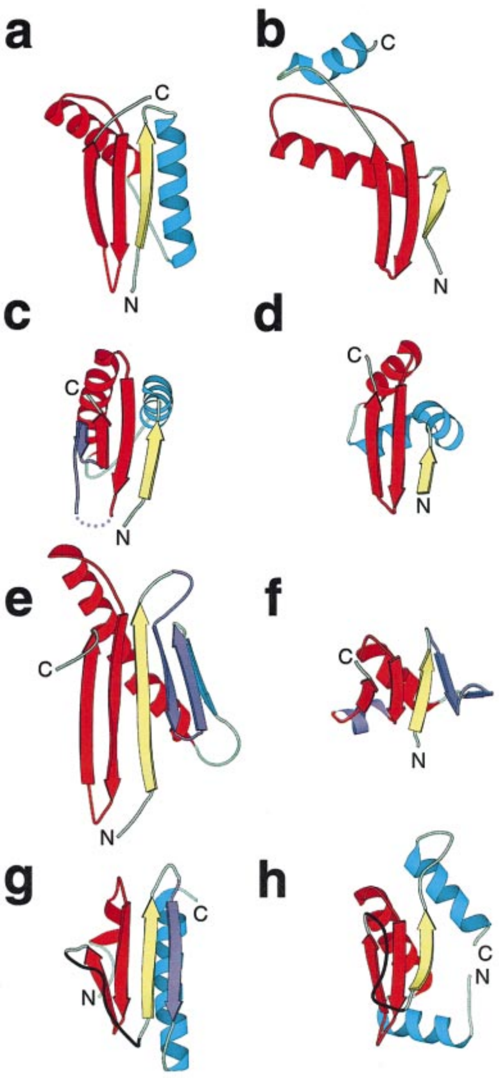


Figure 2. MOLSCRIPT (Kraulis, 1991) ribbon diagrams of protein structures containing the core structural motif of the C-terminal domain of kinase - ATP-grasp fold. The N and C termini of the displayed fragments are labeled in roman upper case black letters. The color-coding scheme corresponds to that described in the legend to Figure 1. β-Strand f is colored yellow, α-helix E and β-hairpin ij are red, α-helices D and F are blue and β-strands i′, g and h are shown in purple. The loop connection between β-strands i and f in circularly permuted structures is colored black. Dots indicate a long insertion that is not displayed. (a) D-Amino acid aminotransferase (Sugio *et al.*, 1988) (1daa, residues A26-A120). (b) Cystatin (Bode *et al.*, 1988) (1cew, residues I9-I87). (c) Replication terminator protein Tus (Kamada *et al.*, 1996) (1ecr, residues A175-A233 and A289-A309). (d) cro λ repressor (Albright & Matthews, 1998) (6cro, residues A2-A57). (e) MHC antigen-recognition domain (Vaughn & Bjorkman, 1998) (3fru, residues A1-A120). (f) Vaccinia DNA topoisomerase I (Sharma *et al.*, 1994) (1vcc, residues 2-63). (g) Villin (Markus *et al.*, 1997) (2vik, residues 1-90). (h) Fe-Mn superoxide dismutase (Ludwig *et al.*, 1991) (3mds, residues A97-A191).

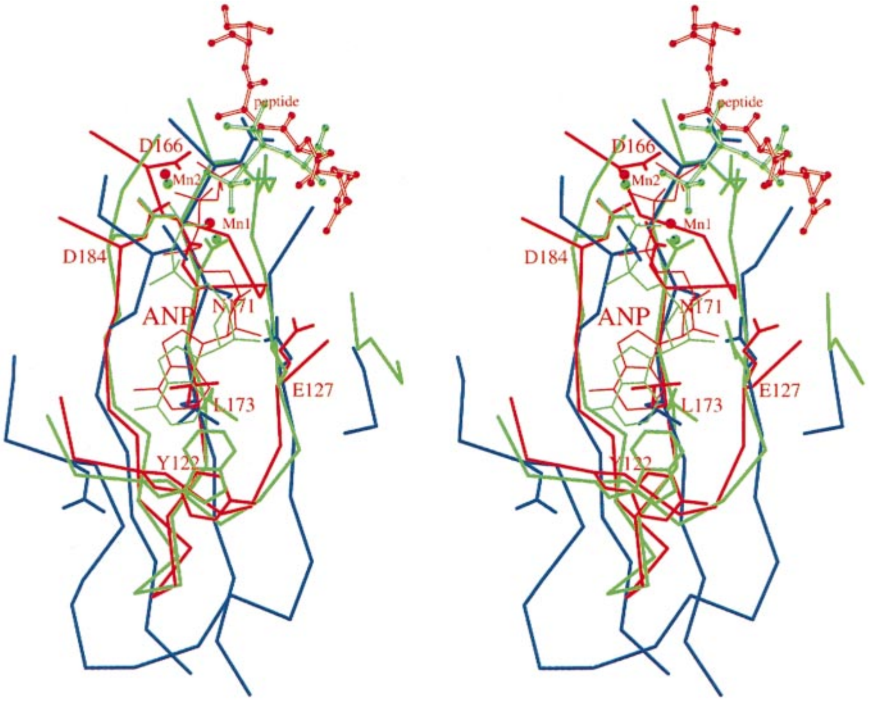


Figure 3. Active site superposition of cAPK (red), PIPK (blue) and DD-ligase (green). Segments of protein chain are shown in thicker lines with the side-chains of functionally important residues displayed. Nucleotides are shown in thinner lines. Inhibitors (peptide for cAPK and phosphinophosphate for DD-ligase) are shown in ball-and-stick presentation. Cations are displayed as balls. Residues are labeled for the cAPK structure. Structurally equivalent residues in other proteins can be deduced from the alignment (Figure 1(e)) or Table 1. ANP stands for AMP-PNP, Mg^{2+} are present in the DD-ligase structure instead of Mn^{2+} in cAPK. The backbone atoms for the following structurally equivalent segments were superimposed to minimize RMSD: cAPK (1cdk), chain B, 121-123, 124-127, 171-176, 177-184; PIPK (1bo1), chain B, 202-204, 213-216, 280-285, 362-369; DD-ligase (2dln), 181-183, 184-187, 257-262, 263-270.

in the active site region. This provides additional evidence for homology between both kinase superfamilies and ATP-grasp proteins.

Although ligands were not considered for the RMSD minimization procedure, the positions of the nucleotide (adenylyl imidodiphosphate (ANP) in cAPK, ADP in DD-ligase) and two metal cations bound to the protein are both essentially identical between cAPK and DD-ligase (Denessiouk *et al*., 1998) (Figure 3 and Table 1A). Conformations of ANP and ADP are strikingly similar, including the orientation of the adenyl ring (*anti* to the ribose with a χ angle of approximately $-135°$), and the 3′-*endo* puckering of ribose and phosphate configuration (Figure 3). Additional argument for homology between these kinase superfamilies and ATP-grasp proteins arises from considerable similarities between their non-nucleotide substrate-binding sites. The enzyme inhibitors (peptide in cAPK and phosphinophosphate in DD-ligase) occupy largely similar structural positions (Figure 3). Moreover, superposition-independent structure-function characteristics, i.e. the distances between equivalent "cofactor" (nucleotide and cation) atoms and interacting protein atoms for cAPK and DD-ligase are correlated (correlation coefficient 0.994) (Table 1B).

Despite the presence of the common substrate/cofactor-binding site, the similarity between PK and ATP-grasp folds is not readily apparent and was not recognized previously (Denessiouk *et al*., 1998). ATP-grasp contains an extended β-sheet and thus is more similar to PIPK than it is to PK. Indeed, comparison of the PIPK and ATP-grasp C-terminal domains reveals the presence of the same major secondary structural elements (f, g, h, E, i, j, F on Figure 1(b) and (d)) with similar arrangement and topological connections. Thus, according to the definition given in SCOP (Hubbard *et al*., 1997), the C-terminal domains of these proteins belong to the same fold. Additionally, consecutive segments of 55 residues comprising four secondary structural elements of PIPK and DD-ligase, namely f, E, i and j, but excluding two regions with gaps, can be superimposed with an RMSD value in C^{α} atom positions of 4.7 Å. The minor differences between the C-terminal domains of DD-ligase and PIPK include a hairpin flip (gh on Figure 1(b) and (d)) and the absence of a short helix D. The helix D, however, is present in some other ATP-grasp proteins, for example in succinyl-CoA synthetase (Wolodko *et al*., 1994) (1scu). Therefore, the structure of PIPK, which shows a clear resemblance to both PK and ATP-grasp, is a link between the two folds.

Table 1. Interatomic distances in cAPK, PIPK and DD-ligase structures

A. *Distances*[a] *between equivalent atoms in the superimposed active sites for CAPK (1cdk), PIPK (1bo1) and* DD*-ligase (2dln)*

	cAPK-DD-ligase	cAPK-PIPK	DD-ligase-PIPK
ANP	1.23 AMP-PNP–ADP	N/A	N/A
Mn1	1.65 Mn–Mg	N/A	N/A
Mn2	1.0 Mn–Mg	N/A	N/A
N171	1.16 N171 $O^{\delta1}$ - D257 $O^{\delta2}$	3.32 N171 $O^{\delta1}$ - S280 O^{γ}	3.78 D257 $O^{\delta2}$ - S280 O^{γ}
D166	4.68 D166 $O^{\delta2}$ - R255 $N^{\eta2}$	4.43 D166 $O^{\delta2}$ - D278 $O^{\delta1}$	4.34 R255 $N^{\eta2}$ - D278 $O^{\delta1}$
D184	1.41 D184 $O^{\delta2}$ - E270 $O^{\varepsilon2}$	4.93 D184 $O^{\delta2}$ - D369 $O^{\delta2}$	3.54 E270 $O^{\varepsilon2}$ - D369 $O^{\delta2}$
	0.31 D184 $O^{\delta1}$ - E270 $O^{\varepsilon1}$	3.04 D184 $O^{\delta1}$ - D369 $O^{\delta1}$	3.28 E270 $O^{\varepsilon1}$ - D369 $O^{\delta1}$
E127	2.63 E127 $O^{\varepsilon2}$ - E187 $O^{\varepsilon1}$	1.98 E127 $O^{\varepsilon2}$ - D216 $O^{\delta2}$	1.94 E187 $O^{\varepsilon1}$ - D216 $O^{\delta2}$

B. *Distances between functionally equivalent atoms (protein - cofactor) for the structures of cAPK (1cdk, in complex with ANP and two* Mn^{2+}*) and* DD*-ligase (2dln, in complex with ADP, PHY and two* Mg^{2+}*)*

cAPK	DD-ligase
2.11 N171 $O^{\delta1}$ - Mn1	1.98 D257 $O^{\delta2}$ - Mg1
3.57 D166 $O^{\delta2}$ - ANP $O^{2\gamma}$	3.28 R255 $N^{\eta2}$ - PHY O^{2P}
1.97 D184 $O^{\delta2}$ - Mn2	1.93 E270 $O^{\varepsilon2}$ - Mg2
2.08 D184 $O^{\delta1}$ - Mn2	2.08 E270 $O^{\varepsilon1}$ - Mg2
2.56 E127 $O^{\varepsilon2}$ - ANP O2′	2.51 E270 $O^{\varepsilon2}$ - ADP O2′

[a] Distances are given in Angstroms, N/A indicates that the comparison is not available due to the absence of nucleotide substrate or cation in the PIPK X-ray structure (Rao *et al.*, 1998). The distance for AMP-PNP to ADP is an RMSD between all equivalent atoms.

Fold similarity does not necessarily indicate homology, since similar folds can originate independently. It is of a special interest, however, to show that the fold similarity between two proteins might be a result of their divergence from the common ancestor, because important functional implications follow homology. It is argued (Murzin, 1998) that not only similarities in the packing of regular secondary structural elements, but conformational similarities in loops and turns may be indicative of homology. For example, the loops formed by residues 120-125 in cAPK and by residues 180-185 in DD-ligase, as well as the turns formed by residues 176-177 in cAPK and by residues 262-263 in DD-ligase, satisfy the latter requirement (Figures 1 and 3).

In summary, the C-terminal domains of PKs, PIPKs and ATP-grasp proteins exhibit an overall structural similarity that manifests itself in an unusual arrangement of regular secondary structural elements combined with the presence of a conserved loop and turn. Additionally, these proteins possess strikingly similar active site structures, including similarities in ligand-binding residues, the conformations of the bound nucleotide, and a substrate-binding site. Taken together, these facts argue that PKs and ATP-grasp proteins are homologs.

The biological implications of this proposal include functional predictions for reaction mechanism and active site residues for PIPK based on what is known of PKs and ATP-grasp proteins. Analysis of the active sites of cAPK and DD-ligase reveals that two Mn^{2+} are coordinated by N166 and D184 in cAPK, and two Mg^{2+} interact with the structurally equivalent D257 and E270 in DD-ligase (Figure 1(a), (d) and (e)). The aliphatic residue (L173 in cAPK and M259 in DD-ligase) is conserved in the adenine-binding pocket of these enzymes. The structure of PIPK complexed with ATP or ATP analog remains to be solved. However, conservation of the catalytic machinery between the PIPKs and PKs, and close superposition of the active site residues of cAPK with their equivalents in PIPK suggests similar functions for these residues (Rao *et al.*, 1998). D369 and S280 might participate in metal binding in PIPK, and L282 should form hydrophobic contacts with the purine ring (Figure 1(b) and (e)). Additionally, D278 in PIPK closely superimposes with D166 in cAPK and might have the same function of a weak base in catalysis (Rao *et al.*, 1998).

PKs, PIPK and ATP-grasp proteins are functionally similar, since they catalyze a phosphotransfer reaction. The main difference is that kinases perform a single-step catalysis, i.e. phosphotransfer needed for signal transduction. The reactions catalyzed by ATP-grasp enzymes are usually more complex and contribute predominantly to macromolecular synthesis. ATP-hydrolysis is used to activate a substrate. For example, DD-ligase transfers phosphate from ATP to D-alanine on the first step of catalysis. On the second step the resulting acylphosphate is attacked by a second D-alanine to produce a DD dipeptide following phosphate elimination (Fan *et al.*, 1994). A similar mechanism was proposed for *Escherichia coli* glutathione synthase (Hara *et al.*, 1996).

Biochemical and structural data support the same "in-line" mechanism of the γ phosphoryl group transfer from ATP to the substrate for PKs and DD-ligase (Ho *et al.*, 1988; Fan *et al.*, 1994). Nevertheless, the remarkable difference between both kinases and DD-ligase is revealed by the substitution of D166 (in cAPK) for R255 (in DD-ligase) (Figure 3). D166 is the closest residue to the γ phosphorus, and was suggested to act as a base for proton abstraction from a substrate nucleophile in PKs (Bossemeyer *et al.*, 1993). Recently, Zhou & Adams (1997) questioned this hypothesis. From the kinetic data they proposed that cAPK D166 does not have the basicity to catalyze proton abstraction, and functions instead as a hydrogen bond acceptor. R255 is unlikely to play either role in DD-ligase. However, the carboxylic group of D-alanine which attacks the γ phosphorus is already negatively charged and does not require to be either an activating base or a H-bond acceptor. Alternatively, positively charged DD-ligase R155 may participate in proper positioning of the D-alanine carboxylate group on the first step of the reaction, and might stabilize the high-energy intermediate for the second step (Fan *et al.*, 1994) (Figure 3). This is an example of how enzymes can change their substrate and reaction specificities merely by amino acid replacements (Asp/Arg) in the key sites.

Comparison of the non-nucleotide substrates of both kinases and ATP-grasp proteins provides an explanation of their structural differences. The extended peptide and protein being the substrates of PKs require an open structure with an extended substrate-binding groove formed by helices D, F, and G (Figure 1). PKs do not have an extended β-sheet structure in the C-terminal domain, which allows additional space for substrate binding. PIPK phosphorylates membrane-embedded phosphoinositides without a requirement for their extraction (Roa *et al.*, 1998; Carpenter & Cantley, 1998). Instead of the groove formed by helices in PK, PIPK possesses a flattened β-sheet near its active site. This flat β-sheet is suggested to be complementary to the surface of the lipid bilayer, which in this case is equivalent to a gigantic substrate. The hairpin gh (Figure 1) shields the active site from the solvent. Most ATP-grasp proteins catalyze two-step reactions and thus require several non-nucleotide substrates. They usually therefore have additional domains that participate in substrate binding (Fan *et al.*, 1994; Wolodko *et al.*, 1994; Hara *et al.*, 1996). Regions of the substrate molecules interacting with the ATP-grasp C-terminal domain are usually small. For example, to facilitate in binding to the small D-alanine molecule, DD-ligase utilizes a helix inserted in the gh hairpin (Fan *et al.*, 1994) (Figure 1(d)). The F-helix, which is positioned much closer to the active site in DD-ligase than in PKs and PIPK, is used for interactions with the second D-alanine molecule. Thus structural rearrangements within the same fold provide a flexibility that would not be achieved by single amino acid substitutions. These rearrangements are so considerable for the case of PKs and ATP-grasp proteins that the structural similarity between them has been very difficult to detect.

Despite the fold similarity in the C-terminal domains of both kinases and ATP-grasp proteins, the N-terminal domains differ in their topologies (Figure 1(a), (b) and (d)), and belong to two distinct but widespread folds†. Recently, the structure of phosphoribosylaminoimidazolesuccinocarboxamide (SAICAR) synthase has been solved (Levdikov *et al.*, 1998) (Figure 1(c)). SAICAR synthase catalyzes the seventh step of purine nucleotide synthesis, namely, condensation of phosphoribosylaminoimidazole with aspartic acid in the presence of ATP and Mg^{2+}. Since this reaction is similar to those catalyzed by ATP-grasp proteins, it is not surprising that the C-terminal domain of SAICAR synthase has the same fold as the C-terminal domain of DD-ligase (Figure 1(c), RMSD 2.25 Å from a superposition of 99 C^{α} atoms; Levdikov *et al.*, 1998). In contrast, the N-terminal domain of SAICAR synthase is not similar in topology to the middle domain of DD-ligase, and instead is similar to the N-terminal domain of PKs (Levdikov *et al.*, 1998; RMSD 2.0 Å in superposition of 38 C^{α} atoms). Therefore SAICAR synthase is topologically similar to PIPK for both its domains (Figure 1(b) and (c)). The N-terminal domain of SAICAR synthase contains a 20-residue insertion into the N-terminal kinase domain fold between β-strands d and e (Figure 1(c), bottom). This insertion results in a reduction of β-strand d in SAICAR synthase when compared with PIPK. The C-terminal domain of SAICAR synthase, in comparison to PIPK, possesses longer β-strands, an additional α-helix between strands f and g, and a flipped hairpin gh. Although the coordinates of SAICAR synthase complexed with ATP are not publicly available, Levdikov *et al.* (1998) demonstrated that the mode of ATP binding in SAICAR synthase is essentially identical with that found in cAPK and DD-ligase, and the structure-based sequence alignment reveals conservation of potential Mg^{2+}-binding and catalytic residues (Figure 1(e)). Thus SAICAR synthase provides another link between PKs and ATP-grasp, and

---

† The N-terminal domain of the ATP-grasph fold is the second ("middle") domain in DD-ligase.

represents an example of an ATP-grasp protein with the N-terminal domain of the kinase fold.

The comparison of PKs and PIPK with other protein structures reveals the structural and functional similarity between the C-terminal domains of kinase and glutathione synthase folds, and suggests that these proteins might be evolutionarily related. The example of SAICAR synthase, which appears to be a hybrid of an N-terminal kinase fold domain and a C-terminal glutathione synthase fold domain, provides additional support for the prediction of homology for PKs and ATP-grasp. The present analysis has been possible only after the determination of the PIPK structure, since this has emphasized the importance of the β-sheet in the kinase fold C-terminal domain. Additionally, the supersecondary structural motif common to the C-terminal domains of PKs, PIPK and ATP-grasp proteins is detected in other protein structures. In contrast to the proposed evolutionary relationship between PKs, PIPK and ATP-grasp proteins argued here, most of these are likely to represent analogs, reflecting the limited number of acceptable spatial arrangements of secondary structural elements.

Acknowledgements

The author thanks Eugene V. Koonin and Chris P. Ponting for critical reading of the manuscript and a number of helpful suggestions, Michael Y. Galperin for various discussions, two anonymous reviewers for helpful suggestions, and Jim H. Hurley for discussions and for making coordinates of PIPK available prior to their release in PDB.

References

Abola, E. E., Sussman, J. L., Prilusky, J. & Manning, N. O. (1997). Protein Data Bank archives of three-dimensional macromolecular structures. *Methods Enzymol.* **277**, 556-571.

Albright, R. A. & Matthews, B. W. (1998). Crystal structure of lambda-Cro bound to a consensus operator at 3.0 Å resolution. *J. Mol. Biol.* **280**, 137-151.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389-3402.

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J. & Turk, V. (1988). The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7**, 2593-2599.

Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H. & Huber, R. (1993). Phosphotransferase and substrate binding mechanism of the cAMP- dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn2+ adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). *EMBO J.* **12**, 849-859.

Carpenter, C. L. & Cantley, L. C. (1998). A flattened face for membranes. *Nature Struct. Biol.* **5**, 843-845.

Daigle, D. M., McKay, G. A., Thompson, P. R. & Wright, G. D. (1999). Aminoglycoside antibiotic phosphotransferases are also serine protein kinases. *Chem. Biol.* **6**, 11-18.

Denessiouk, K. A., Lehtonen, J. V., Korpela, T. & Johnson, M. S. (1998). Two "unrelated" families of ATP-dependent enzymes share extensive structural similarities about their cofactor binding sites. *Protein Sci.* **7**, 1136-1146.

Fan, C., Moews, P. C., Walsh, C. T. & Knox, J. R. (1994). Vancomycin resistance: structure of D-alanine:D-alanine ligase at 2.3 A resolution. *Science*, **266**, 439-443.

Hara, T., Kato, H., Katsube, Y. & Oda, J. (1996). A pseudo-michaelis quaternary complex in the reverse reaction of a ligase: structure of *Escherichia coli* B glutathione synthetase complexed with ADP, glutathione, and sulfate at 2.0 Å resolution. *Biochemistry*, **35**, 11967-11974.

Ho, M.-F., Bramson, N. H., Hansen, E. D., Knowles, R. J. & Kaiser, E. T. (1988). Stereochemical course of the phospho group transfer catalyzed by CAMP-depenant protein-kinese. *J. Am. Chem. Soc.* **110**, 2680-2681.

Holm, L. & Sander, C. (1997). Dali/FSSP classification of three-dimensional protein folds. *Nucl. Acids Res.* **25**, 231-234.

Hon, W. C., McKay, G. A., Thompson, P. R., Sweet, R. M., Yang, D. S., Wright, G. D. & Berghuis, A. M. (1997). Structure of an enzyme required for aminoglycoside antibiotic resistance reveals homology to eukaryotic protein kinases. *Cell*, **89**, 887-895.

Hubbard, T. J. P., Murzin, A. G., Brenner, S. E. & Chothia, C. (1997). SCOP: a structural classification of proteins database. *Nucl. Acids Res.* **25**, 236-239.

Kajava, A. V. (1992). Left-handed topology of super-secondary structure formed by aligned alpha-helix and beta-hairpin. *FEBS Letters*, **302**, 8-10.

Kamada, K., Horiuchi, T., Ohsumi, K., Shimamoto, N. & Morikawa, K. (1996). Structure of a replication-terminator protein complexed with DNA. *Nature*, **383**, 598-603.

Knighton, D. R., Zheng, J. H., Ten, Eyck L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S. & Sowadski, J. M. (1991a). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*, **253**, 407-414.

Knighton, D. R., Zheng, J. H., Ten, Eyck L. F., Xuong, N. H., Taylor, S. S. & Sowadski, J. M. (1991b). Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*, **253**, 414-420.

Kobayashi, N. & Go, N. (1997). ATP binding proteins with different folds share a common ATP-binding structural motif. *Nature Struct. Biol.* **4**, 6-7.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Levdikov, V. M., Barynin, V. V., Grebenko, A. I., Melik-Adamyan, W. R., Lamzin, V. S. & Wilson, K. S. (1998). The structure of SAICAR synthase: an enzyme in the de novo pathway of purine nucleotide biosynthesis. *Structure*, **6**, 363-376.

Ludwig, M. L., Metzger, A. L., Pattridge, K. A. & Stallings, W. C. (1991). Manganese superoxide dismutase from *Thermus thermophilus*. A structural model refined at 1.8 Å resolution. *J. Mol. Biol.* **219**, 335-358.

Markus, M. A., Matsudaira, P. & Wagner, G. (1997). Refined structure of villin 14T and a detailed comparison with other actin-severing domains. *Protein Sci.* **6**, 1197-1209.

Murzin, A. G. (1998). How far divergent evolution goes in proteins. *Curr. Opin. Struct. Biol.* **8**, 380-387.

Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. & Thornton, J. M. (1997). CATH-a hierarchic classification of protein domain structures. *Structure*, **5**, 1093-1108.

Rao, V. D., Misra, S., Boronenkov, I. V., Anderson, R. A. & Hurley, J. H. (1998). Structure of type IIbeta phosphatidylinositol phosphate kinase: a protein kinase fold flattened for interfacial phosphorylation. *Cell*, **94**, 829-839.

Sharma, A., Hanai, R. & Mondragon, A. (1994). Crystal structure of the amino-terminal fragment of vaccinia virus DNA topoisomerase I at 1.6 Å resolution. *Structure*, **2**, 767-777.

Sugio, S., Petsko, G. A., Manning, J. M., Soda, K. & Ringe, D. (1995). Crystal structure of a D-amino acid aminotransferase: how the protein controls stereoselectivity. *Biochemistry*, **34**, 9661-9669.

Vaughn, D. E. & Bjorkman, P. J. (1998). Structural basis of pH-dependent antibody binding by the neonatal Fc receptor. *Structure*, **6**, 63-73.

Wolodko, W. T., Fraser, M. E., James, M. N. & Bridger, W. A. (1994). The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5-A resolution. *J. Biol. Chem.* **269**, 10883-10890.

Zhou, J. & Adams, J. A. (1997). Is there a catalytic base in the active site of cAMP-dependent protein kinase? *Biochemistry*, **36**, 2977-2984.

Edited by J. M. Thornton

(Received 24 March 1999; received in revised form 24 June 1999; accepted 28 June 1999)