Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochimica et Biophysica Acta xx (2005) xxx – xxx
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
http://www.elsevier.com/locate/bba

Review

The RIO kinases: An atypical protein kinase family required for ribosome biogenesis and cell cycle progression☆

Nicole LaRonde-LeBlanc, Alexander Wlodawer*

Protein Structure Section, Macromolecular Crystallography Laboratory, NCI-Frederick, Bldg. 536, Rm. 5, Frederick, MD 21702-1201, USA

Received 9 June 2005; received in revised form 6 July 2005; accepted 8 July 2005

Abstract

Atypical protein kinases (aPKs) include proteins known to be involved in the phosphorylation-mediated regulation of a wide variety of cellular processes, as well as some for which the function is, as yet, unknown. At present, 13 families of aPKs have been identified in the human genome. This review briefly summarizes their known properties, but concentrates in particular on the RIO family of aPKs. Representatives of this family are present in organisms varying from archaea to humans. All these organisms contain at least two RIO proteins, Rio1 and Rio2, but a third Rio3 group is present in multicellular eukaryotes. Crystal structures of *A. fulgidus* Rio1 and Rio2 have shown that whereas the overall fold of these enzymes resembles typical protein kinases, some of the kinase structural domains, particularly those involved in peptide substrate binding, are not present. The mode of binding of nucleotides also differs from other kinases. While the enzymatic activity of Rio1 and Rio2 has been demonstrated and both have been shown to be essential in *S. cerevisiae* and required for proper cell cycle progression and chromosome maintenance, the biological substrates of RIO proteins still remain to be identified.

Keywords: Atypical protein kinase; Structure; ATP binding; Enzymatic activity; Ribosome biogenesis

1. Introduction

The protein kinases are critical for regulating a large variety of cellular processes and thus many of them are becoming extremely important drug targets. In human cells, 518 protein kinases that have been identified so far catalyze transfer of phosphates to serine, threonine and tyrosine residues [1]. These enzymes, structurally exemplified by the cyclic adenosine monophosphate-dependent protein kinase (PKA) [2,3], are characterized by the presence of a catalytic domain of 250 to 300 amino acids (solely or in combination with regulatory domains) which contains conserved residues that play a role in nucleotide binding, peptide substrate binding, and phosphoryl transfer. The classical protein kinase fold consists of an N-terminal lobe containing a β-sheet adjacent to a single α-helix (αC) and a C-terminal lobe that is mostly helical. These lobes are connected by a short flexible linker that allows movement of the two lobes relative to each other when nucleotide binds in the cleft between them. The conserved residues are located within functional domains, or subdomains, that are used to describe structural details of protein kinases [4]. Among these subdomains are a nucleotide-binding loop (subdomain I), typically with the sequence GXGXXG, which binds and orients the phosphates of ATP; a hinge region which interacts with the adenine moiety of the ATP via hydrogen bonds and hydrophobic interactions; a catalytic loop (subdomain VIb) which contains conserved catalytic Asn and Asp residues directly involved in phosphoryl transfer; and a metal-binding or "DFG" loop (subdomain VII) with a conserved Asp required for the positioning of metal ions. Canonical eukaryotic protein kinases (ePKs) also contain a loop between the metal-binding and the catalytic loop known as the "APE" or activation loop (subdomain VIII). In these kinases, phosphorylation of this loop results in modulation of the kinase activity [5,6]. The activation loop is also critical for binding and recognition of a peptide substrate [3,5]. Additional peptide substrate binding surface is provided by C-terminal helices known as subdomains X and XI in the structure of PKA bound to an inhibitory peptide, PKI [3].

☆ Dedicated to Professor David Shugar on the occasion of his 90th birthday.

* Corresponding author. Tel.: +1 301 846 5036; fax: +1 301 846 6322.
E-mail address: wlodawer@ncifcrf.gov (A. Wlodawer).

1570-9639/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.bbapap.2005.07.037

2. Atypical protein kinases

Recent analysis of the set of protein kinases in the human genome, termed the "kinome", has revealed several proteins with confirmed protein kinase activity which have little sequence similarity to any known eukaryotic protein kinases (ePKs) [1]. These kinases are called atypical protein kinases (aPKs) and include proteins known to be involved in the phosphorylation-mediated regulation of a wide variety of cellular processes, as well as some for which the function is yet unknown. Atypical protein kinases should not be confused with atypical protein kinase C's, which are different from PKC (calcium-dependent protein kinase) in the regulatory regions of the molecules, but retain homology to PKC in the kinase domain. Of the 518 identified human kinases, 40 are classified as atypical. These 40 fall into 13 families or homology groups. Unlike ePKs, each of the identified aPK families is represented in the human kinome by only a few members (2–6). Many of these aPKs have been shown to bear significant structural homology to ePKs, despite the lack of sequence similarity, while others are structurally distinct [1]. Some of the groups are restricted to metazoans, while others show conservation even in prokaryotes. A brief description of what is known to date about each of the 13 families identified by Manning et al. is included below. Table 1 provides a brief summary of the functional information available and, if known, the amino acid(s) that they phosporylate. Subsequently, we will discuss in significantly more detail the properties of the RIO family of aPKs.

2.1. Alpha kinases

The α-kinases form a family which includes EF-2 kinase, a molecule shown to phosphorylate elongation factor-2 [7,8]. This family also includes the myosin heavy chain kinases of *Dictyostelium discoideum*, as well as the ~ 300 residue domain of ChaK [9,10]. The crystal structure of the kinase domain of ChaK (channel kinase) is the only one available for an α-kinase [11]. Despite the lack of sequence homology with the known protein kinases, the structure of ChaK kinase domain is homologous to ePK kinase fold. It consists of two globular domains, the N-terminal mostly β-sheet lobe and the C-terminal mostly α-helical lobe, connected by a flexible linker (Fig. 1A). As seen in canonical protein kinases, the ATP molecule binds in the cleft between the two lobes. The activation loop is highly conserved among the α-kinases and contains a glycine-rich sequence, which is thought to participate in substrate interaction. Although there is significant similarity between the α-kinase catalytic domain and the ePK kinase domain, the C-terminal lobe of the α-kinase domain contains several distinct features. The C-terminal lobe of ChaK contains a zinc-binding module required for structural stability of the domain, unlike ePKs, and a Gln residue located two positions away from the catalytic Asp replaces the catalytic Asn residue which is located five positions away from the catalytic Asp in ePKs [11]. The ChaK kinase domain has been shown to phosphorylate myelin basic protein (MBP) on both serine and threonine residues [10].

Table 1
Atypical protein kinase families

Kinase	Type	Functional information
A6 K	Tyr	unknown
ABC1	ND	unknown
Alpha	Ser/Thr	Translation regulation (EF2 Kinase); Ion channel kinase (TRP-PLIK/ChaK).
BCR	Ser/Thr	Fusion partner in Bcr-Abl; downregulates Ras signaling by phosphorylating AF-6 and 14–3–3.
BRD	ND	Bromodomain containing; transcriptional regulators a.k.a BET proteins; meiosis, cell cycle control, homeosis. Eg: BRD2/Ring3, MCAP/BRD4.
FAST	Ser/Thr	Apoptosis; activated downstream of the Fas antigen.
G11	Ser/Thr	Located in the major histocompatibility locus.
H11	Ser/Thr	Heat shock protein (H11/HspB8); contains α-crystallin domain.
PDK	Ser	Regulation of oxidation of pyruvate (PDK), or branched chain α-ketoacids (BCKD). Both are mitochondrial.
PIKK	Ser/Thr	Stress response (DNA-PK, ATM, ATR, SMG-1), translation regulation (m-TOR).
RIO	Ser	Ribosome biogenesis (Rio1, Rio2), cell-cycle progression (Rio1).
TAF	Ser/Thr	Transcription initiation (TAF II-250/TFIID).
TIF	Ser/Thr	Transcription regulation (TIF1-α).

2.2. The A6 kinases

The founding members of this family are the human A6 and A6r gene products [12,13]. A6 kinase, also known as PTK9, was shown to exhibit tyrosine kinase activity in vitro when produced as fusion proteins in bacteria. However, subsequent studies have shown that these proteins interact with PKCζ and bind ATP, but did not detect kinase activity [13]. Therefore, the inferred kinase activity of this group is based on a single report [12]. A6 kinases do not show significant sequence homology to known ePKs and their structure is at present unknown.

2.3. The phosphoinositide 3′ kinase-related kinases (PIKK)

This family contains large proteins such as mTOR (mammalian target of rapamycin), DNA-PK (DNA-dependent protein kinase), and ATM (ataxia telangiectasia mutated) [14]. Members of this family contain a kinase domain similar to the phosphoinositide 3′ kinase (PI3K) domain (Fig. 1B). Despite significant similarity to these lipid kinases, members of the PIKK family only phosphorylate proteins. The PI3K domain is indeed structurally related to the ePK kinase domain, containing two lobes joined by a linker that binds ATP in the cleft in between them [15]. The significant difference between the PI3K domain and the ePK kinase domain is that the loop in PI3K that is homologous to the P-loop of ePKs contains no glycine, and instead contacts the triphosphate group via a side chain interaction from a conserved serine in the loop. PI3K catalytic domains also contain an activation loop segment analogous to that seen in ePKs, but with a distinct sequence which determines the sequence specificity of the PI3Ks.

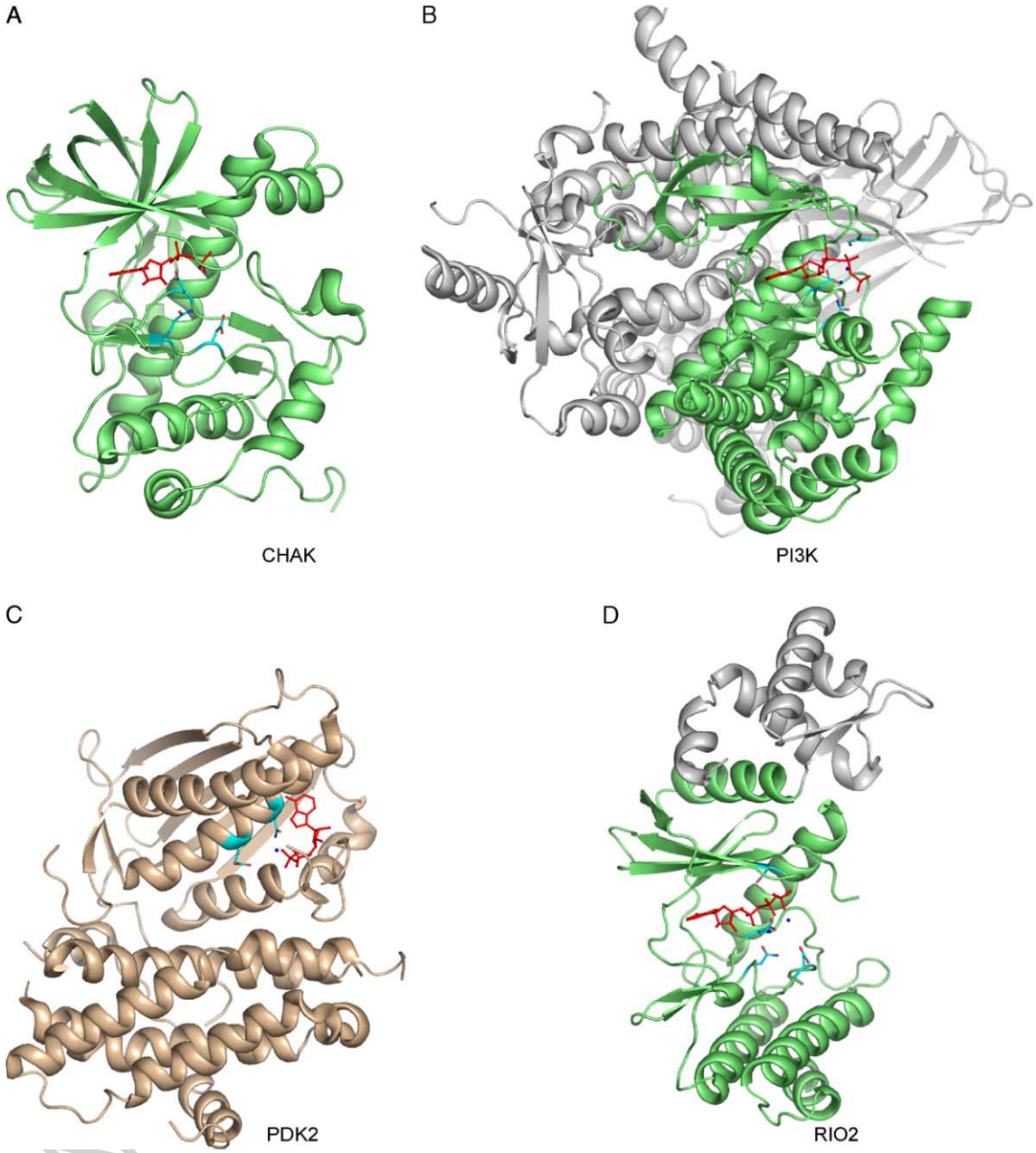


Fig. 1. Structures of the kinase domains of atypical protein kinases. (A) ChaK (PDB code: 1IA9); (B) PI3K (phosphoinositol 3' kinase; PDB code: 1E8X); (C) PDK2 (PDB code: 1JM6); (D) Rio2 (PDB code: 1ZAO). The green-colored domains are homologous to the canonical ePK kinase domain. The kinase domain of PI3K is expected to be similar to that of the PIKK group of atypical protein kinases. The catalytic domain of wheat PDK2 is homologous to histidine kinases. The catalytic residues are highlighted in cyan, and ATP or ATP analogue is shown bound to the active site of the molecules.

2.4. The ABC1 kinases

The founding member of this family of kinases is ABC1 from yeast and includes AarF from *E. coli*. The ABC1 kinases are not related to the family of ATP transport proteins. In yeast the protein is mitochondrial and necessary for coenzyme Q synthesis. Leonard et al. [16] noted that the proteins of the ABC1 family, of which there are representatives in a diverse range of organisms from bacteria to humans, contain a kinase signature with conserved important catalytic and metal binding residues of the ePK kinase domain. The exact function of these proteins in mammals has not been defined, and there is no structure of any of them.

2.5. The bromodomain kinases (BRD)

This family was first discovered through characterization of a nuclear kinase named RING3 or BRD2 [17]. The founding members of the family are human BRD2, fruit fly Fsh, and yeast Bdf1p [18]. BRD2 is known to drive leukemogenesis when overexpressed in mice [19]. These proteins contain a conserved region that shows weak homology to the ePK kinase domain, and contains two bromodomains. The BRD2

protein and its homologs are thought to be transcription factor kinases and BRD2 has been shown to interact with E2F, a cell-cycle regulating transcription factor, in the presence of acetylated histones [20]. In complexes with E2F, BRD2 transactivates the promoters of E2F-dependent cell cycle genes [20]. The structure of the kinase domain of these proteins is still unknown.

2.6. The BCR kinases

The BCR (breakpoint cluster region) kinase is known as a component of BCR-Abl, the fusion protein of BCR with the Abl nonreceptor tyrosine kinase which is present in up to 95% of cases of chronic myeloid leukemia [21]. The unfused BCR is a large (145 kDa) protein containing several domains, including an oligomerization domain, a DH/PH pair of domains (guanine nucleotide exchange and pleckstrin homology) and a Rho-GAP domain [22]. BCR is known to form tetramers on its own, or as part of the BCR-Abl fusion protein. In 1991, it was reported that purified BCR contained autophosphorylation activity and could phosphorylate other substrates [23]. Since then, additional substrates for the kinase have been identified [24]. The kinase domain was mapped to within the N-terminal 400 residues of the protein [23]. This region does not show any relevant homology to known kinase domains and the structure of the kinase domain is unknown.

2.7. The H11 kinase

The human H11 gene was identified in a search for human homologs of the ICP10 protein kinase (ICP10PK) of the herpes simplex virus [25,26]. H11 is overexpressed in melanoma cells and is 30% identical to ICP10PK and also displays Mn^{2+}-dependent Ser/Thr autophosphorylation activity which is blocked by mutation of the putative ATP binding Lys residue from subdomain II [25]. Later, it was noted that H11 belongs to the family of small heat shock proteins characterized by a conserved α-crystallin domain and the protein was renamed HSPB8 [27]. Although the HSP20 family contains several other members, kinase activity has not been demonstrated for any others besides H11. Only very speculative conservation of ePK subdomains was noted and the structure of H11 remains unknown.

2.8. The Fas-activated s/t kinases (FASTK)

The FAST kinase was identified in a screen aimed at identifying proteins which bind to TIA-1, an RNA-binding protein which is an effector of apoptotic cell death [28]. Sequence analysis revealed that it contained limited sequence similarity to the ICP10 protein kinase domain of the herpes simplex virus. FASTK was shown to have serine/threonine autophosphorylation activity, and phosphorylation activity on TIA-1. Weak sequence similarity to ePKs was noted, although clear candidates for the catalytic residues were not identified.

2.9. The pyruvate dehydrogenase kinases (PDK)

These kinases are mitochondrial and specifically phosphorylate the E1 subunit of the pyruvate dehydrogenase complex, thereby regulating the activity of the complex and the flow of energy from glycolysis to oxidation or storage [29]. Activation or induction of PDKs results in inactivation of E1 by phosphorylation, which corresponds to an increase in serum glucose levels as seen in diabetic hyperglycemia [30]. These kinases, for which four isozymes have been identified in humans, have no homology to ePKs but have significant homology to bacterial histidine kinases. PDK2, which is the most abundantly expressed isozyme, phosphorylates only serine residues. Structures were determined for the PDK2 (Fig. 1C) and a related enzyme, BCK (branched-chain α-ketoacid dehydrogenase kinase), from rat [31]. The structures showed that these kinases are indeed structurally homologous to bacterial histidine kinases, and in both cases formed a dimer in the crystal which was also confirmed in solution.

2.10. TATA binding factor associated factor 1 (TAF1)

In 1996, Dikstein et al. [32] reported that TAF1, also known as TAF II-250, is a protein kinase. TAF1 is a part of the transcription initiation factor TFIID and plays a role in basal transcription initiation. TAF1 was shown to contain two kinase domains which conserve the catalytic residues of ePK domains. Kinase activity was demonstrated for both domains, and the protein was shown to specifically phosphorylate RAP74, a component of TFIIF [32]. Other than the conserved catalytic residues, very little sequence homology was noted between TAF1 and protein kinase domains. Later studies confirmed the kinase activity of TAF1, and established that TFIIA, another subunit of the transcription complex, is phosphorylated by TAF1 [33]. TAF1 also contains two bromodomains for which the crystal structure was solved, but the structure of the kinase domain remains unknown.

2.11. Transcription intermediary factor 1 (TIF1)

This family contains three related proteins, TIF1α, β, and γ, involved in regulation of the transcription machinery. TIF1α was shown to have kinase activity, and due to the high level of conservation (43% similarity with TIF1β and 77% similarity with TIF1γ), the other TIF1 proteins are expected to have kinase activity as well [34]. Autophosphorylation activity was detected, and the ability to phosphorylate TFIIEα, TAFII28, and TAFII55 in vivo was also reported. These proteins contain an RBCC (RING finger-B boxes-coiled coil) motif on the N-terminal end and a PHD finger and a bromodomain on the C-terminal end. The location of the kinase motif has not been identified.

3. RIO kinases: ancient molecules linked to kinase evolution

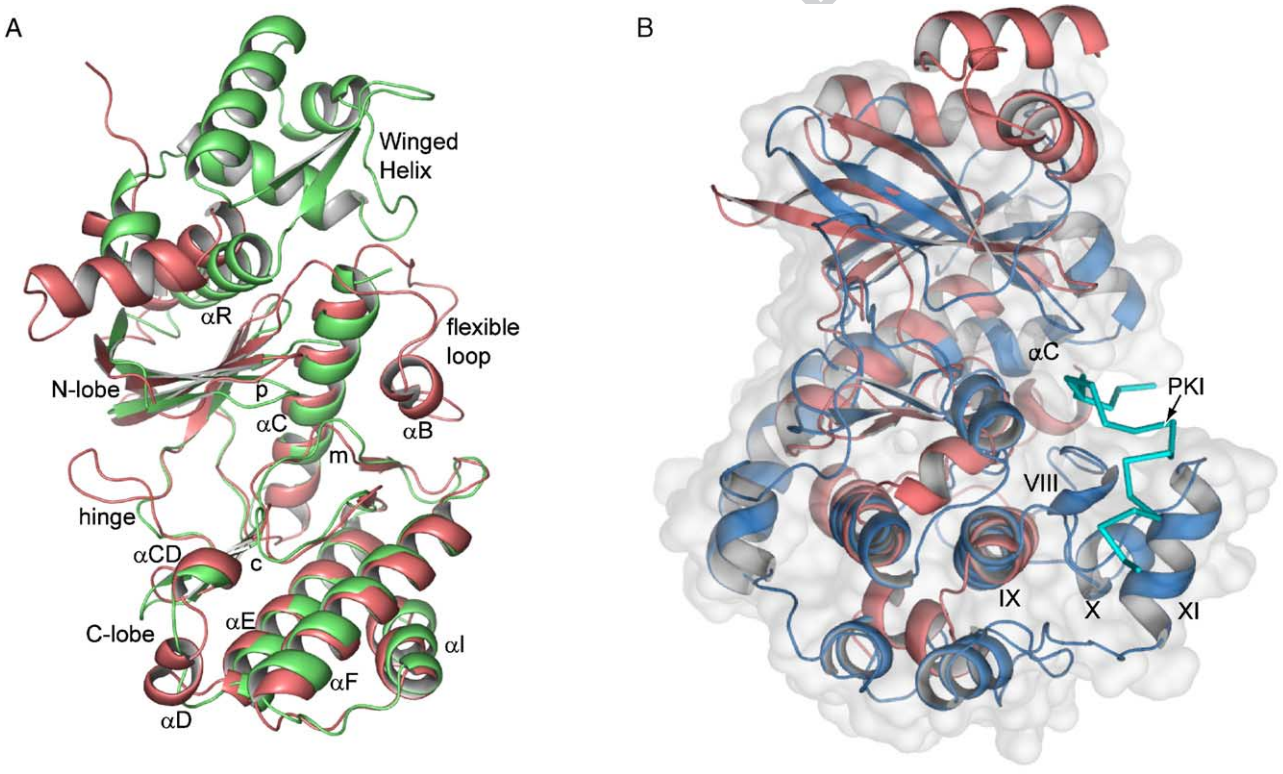
The RIO family was first identified as a group of proteins containing the conserved RIO domain, named based on the founding member of the family, yeast Rio1 (right open reading

frame). The RIO domain contains a discernible kinase signature, but otherwise exhibits little sequence similarity with ePKs [35]. Representatives of this family are present in organisms varying from archaea to humans. All these organisms contain at least two RIO proteins, one which is more similar to yeast Rio1, and one with a moderately different RIO domain and a conserved N-terminal domain, homologous to yeast Rio2. Eventually, a third group of RIO proteins, designated Rio3, was discovered. Members of the Rio3 subfamily, which is more similar to Rio1, but also contains a conserved N-terminal domain different than that of Rio2, have been found thus far only in multicellular eukaryotes. As shown in Fig. 2A and B, each RIO subfamily is distinguished by specific sequence variations in the RIO kinase domains, as well as the presence of subfamily-specific, conserved N-terminal sequences.

It was also reported that a group of bacterial kinases bear significant sequence homology to the RIO kinases. These bacterial RIO kinases are present in a few species of bacteria, and there is only a single representative per organism. Examination of the sequences revealed that the bacterial RIO kinases are more similar to Rio1 in the N-terminal half, and more similar to Rio2 in the C-terminal half of the kinase domain. Thus, it appears that the bacterial RIO kinases are related to both enzymes, and may represent the remnants of a common progenitor of the two subfamilies. This is interesting because it has also been reported that the KDO lipid kinases bear significant homology to the bacterial RIO kinases, and thus the RIO kinases may represent the evolutionary link between bacterial lipid kinases and ePKs [16,36].

3.1. RIO kinases and ribosome biogenesis

The founding member of the RIO family, Rio1, is an essential gene in *S. cerevisiae*, required for proper cell cycle progression and chromosome maintenance [35]. In yeast cells



C

Kinase	I	II	III	V	VIB	VII	IX
Rio2	xxoGVGKESxoYxx	xoKoHRxGxTSFRxoKxK	oSxxxRxxAx+EFxooxxL	HxVoMxLo-----xGx	GoIHxDoNEFNooo	oIDFPQxxSxxHx	oxRDV
Rio1	GxISTGKEANVYHA	AIKIYKTSxxxFKDx-+Y	PRKxVxxWAEKExRNLxxR	HVLVMxFxxxDxxPAP	+LVHADLSEYNxoo	oIDVSQSVxxxHP	LRxDo
Rio3	GxISTGKESVVFHA	AIKIFKTTLNEFKNRDKY	PRKIxRMWAEKEMHNLxRM	HoLVMxFIG+DQxPAP	xLVHADLSEYNMLW	oIDVSQSVEPTHP	LoRDC
RioB	xxLxSGKEAxVYxx	xxKVYKExxxRSFKxxAX	xxxxExxWxxAEoxxLoxL	xVLLMxLVxxxxxxAP	GoVHGDLSEFNVLo	IIDoPQxVxxAxx	oxRDo
PKA	KTLGTGSFGRVMLV	AMKILDKQKVVKLK----	----QIEHTLNEKRILQAV	LYMVMEYVxxxxxPGG	DLIYRDLKPENLLI	VTDFGFAKRVKGR	KAVDW
CK1	RRIGEGSFGVIFEG	AIKFEPRRSDAPQ-----	--------LRDEYYRTYKL	NVLVIDLLxxxxxGPS	SLVYRDIKPDNFLI	VVDFGMVKFYRDP	RRDDL

Fig. 2. Structural features of the RIO kinases. (A) The structure of the Rio1 kinase (red; PDB code: 1ZTF) superimposed on that of Rio2 (green; PDB code: 1TQI). (B) The structure of the ATP-bound form of Rio1 (red; PDB code: 1ZP9) superimposed on that of PKA (blue; PDB code: 1ATP). A transparent surface representation of the PKA and a backbone representation of the bound peptide inhibitor PKI (cyan) are shown. Roman numerals indicate subdomains. (C) Alignment of the conserved sequences of the four RIO subfamilies with the sequences of PKA and casein kinase (CK1). Red, green, and blue text represents identical, highly similar or weakly similar positions as determined by ClustalW alignment of several representatives of each group. The symbols o and + represents positions where hydrophobic or charged residues, respectively, are conserved.

deprived of Rio1, cell cycle arrest occurs in G1 or mitosis, indicating Rio1 activity is required for entry into S phase and exit from mitosis [35]. In addition, Rio1 and Rio2 were identified as non-ribosomal factors necessary for late 18S rRNA processing. In yeast, depletion of Rio1 or Rio2 also affects growth rate and results in an accumulation of 20S rRNA [37–39]. Deletion of either Rio1 or Rio2 is lethal, suggesting that the two proteins perform distinct functions [38,40]. It has been demonstrated that the yeast RIO proteins are indeed capable of serine phosphorylation in vitro, and conserved kinase catalytic residues are required for their in vivo function [35,37,39].

3.2. Structural characteristics of the RIO domain

The RIO kinase domain was first characterized structurally on the basis of the crystal structure of the full-length *Archaeoglobus fulgidus* Rio2 protein (Fig. 1D) [41]. The RIO domain is structurally homologous to kinase domains but is surprisingly small, truncated by deletion of the loops known to be important for substrate binding in ePKs (subdomains VIII, X, and XI). This is unexpected since subdomain VIII, also known as the "activation loop", was thought to be absolutely necessary in order to provide peptide recognition and binding and thus enable protein kinase activity. Analysis of the amino acid sequences indicates that the absence of this loop appears to be a feature of all RIO kinases, including their eukaryotic versions. The RIO domain contains the β-sheet N-lobe and the α-helical C-lobe connected by a flexible hinge region, as seen in typical protein kinases. When the sequence of the RIO domain is aligned with the kinase domain of PKA, the catalytic loop and metal binding loop residues are in a similar position. Comparison with the subsequently determined structure of the Rio1 protein [42] from the same organism revealed that the minimal RIO domain also includes a helix N-terminal to the canonical N-lobe, and a loop inserted between the third β strand of the N-lobe and the α helix C (Fig. 2C). The sequences of this insertion are conserved only within each subfamily and the residues in this region form a small helix packed against the side of the molecule near the active site in Rio1, but are largely disordered in the structure of Rio2.

3.3. Nucleotide binding by the RIO kinases

Although several of the key residues involved in catalysis in the typical kinase domain are conserved in the RIO kinases, several differences exist in the active site, corresponding to differences in how RIO kinases interact with ATP. The canonical phosphate-binding loop (or P-loop), as seen in PKA, contains several glycines (GxGxxG) and the lack of the side chains facilitates direct interactions between the phosphate groups of the ATP and the backbone of the P-loop. However, the RIO kinases have subfamily-specific loops, with the sequence STGKEA for Rio1, GXGKES for Rio2, and STGKES for Rio3, significantly different than their counterparts in ePKs. This results in significant differences in how the RIO proteins interact with the phosphates. In Rio1, direct contacts are made from the side chain of the invariant Ser in the start of the P-loop, to the β phosphate (Fig. 3A). In Rio2, this contact is replaced by the invariant Ser in at the end of its P-loop sequence (Fig. 3B). In the case of both Rio1 and Rio2, the phosphate is bound in an extended conformation that is significantly altered from that seen in most active protein kinase-ATP complexes (Fig. 3C). The conformation of the phosphate bound to Rio3 remains to be established since its P-loop sequence contains two conserved serine residues.

The coordination of a metal ion between the α and β phosphate is observed in the structures of Rio1 and Rio2 (Fig. 3). In Rio2, an additional metal ion is seen between the γ and β phosphates (Fig. 3B). The observation of a single metal in Rio1 versus two metal ions in Rio2 may be due to low occupancy of the second site that results from partial hydrolysis of the γ-phosphate in Rio1. Much remains to be elucidated regarding the occupancy of metal ion sites upon peptide substrate binding. The metal ion that is seen in the structures of both Rio1 and Rio2 is in fact conserved in other protein kinases. However, in those kinases, such as PKA, the metal ion is coordinated by the α- and γ-phosphates, and plays a role in catalysis through direct interaction with the leaving phosphate group (Fig. 3C). The importance of this discrepancy between RIO kinases and canonical ePKs remains to be explored.

A comparison of the ATP binding pocket of the RIO kinases with that of the ePKs has revealed several unique features. Fig. 4 illustrates the differences in the ATP binding cavity of Rio1, Rio2, and PKA and indicates differences between the RIO kinases and canonical ePKs in terms of the charge distribution in the active site, and the placement of cavities. The figure also shows significant differences between the Rio1 and Rio2 cavities, illustrating more open access to the γ-phosphate in Rio1, and more open access to the adenine moiety in Rio2. The environment surrounding the ribose moiety is also unique for each enzyme. This is important in considering the design of inhibitors that will not only be RIO kinase specific, but RIO subfamily specific as well.

3.4. The flexible loop of RIO domains

A region between the third β sheet of the RIO domain N-lobe and αC is disordered in the structure of Rio2 and was therefore called the flexible loop. This region is 18-residue long in Rio2 and 27-residue long in Rio1. In the structure of Rio1 without bound ATP, the entire region was traceable in the electron density, but in the presence of ATP and ADP, small portions near the ends of this region were not seen. This observation illustrates that a high degree of flexibility is exhibited by this loop in Rio1 as well. In Rio1, the flexible loop forms a small α helix which binds to the side of the molecule via hydrophobic and hydrophilic interactions. The position of this helix relative to the rest of the molecule is altered depending upon the presence of the triphosphate group, representing a large conformational change which occurs in the molecule in response to nucleotide binding. As such, this part of the molecule may participate in regulation of the activity of

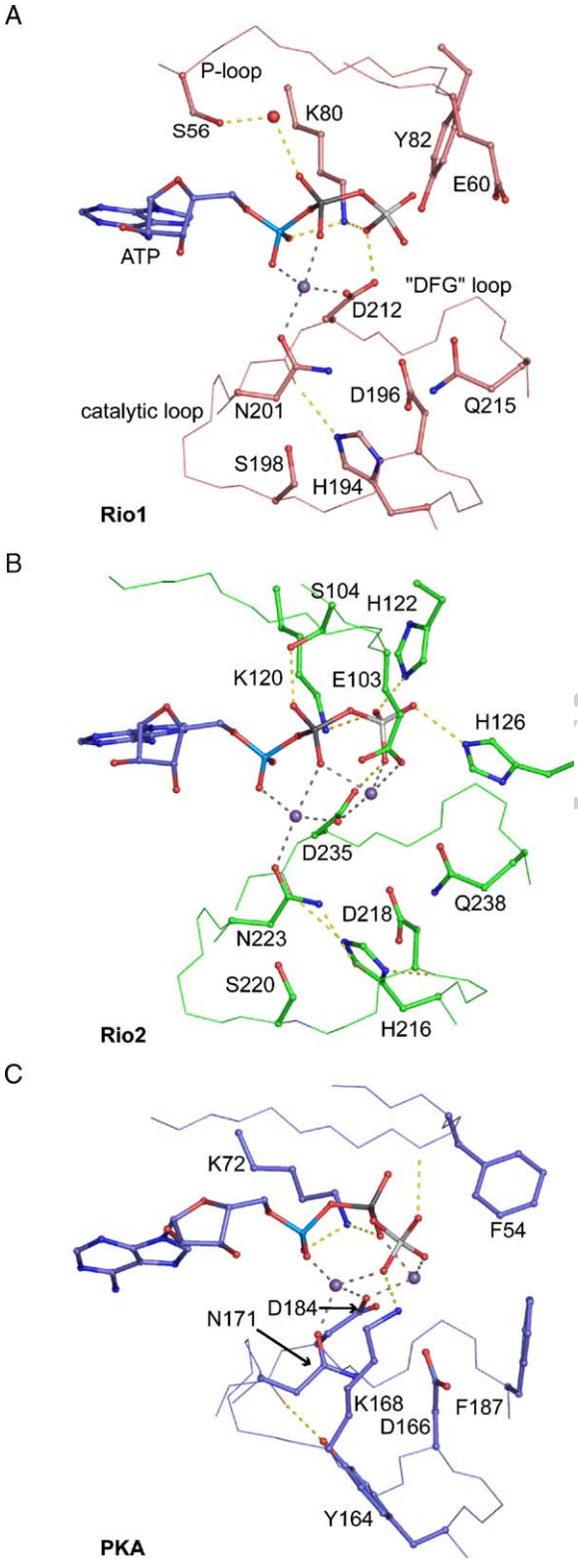


Fig. 3. The active sites of the RIO kinases. The structure of the active sites of (A) Rio1, (B) Rio2 and (C) PKA are shown with bound ATP and metal ions (purple spheres). The backbone of the phosphate-binding loop (P-loop), the metal binding loop ("DFG" loop), and the catalytic loop are shown, with the catalytic and phosphate-binding residues in stick representation.

the molecule, and the positioning of side chains originating from the flexible loop may influence the position of residues that directly interact with the triphosphate moiety.

The sequence of the flexible loop (subdomain V) is conserved among subfamily members but not between them (Fig. 2B). All three identified subfamilies contain a specific sequence in this region. Therefore, the function of this region may be different for each RIO kinase in an organism, and may play a role in the subfamily specific functions of these kinases. The family of bacterial RIO kinases contains a conserved sequence in this region as well, and this sequence is different from all of the other RIO kinases. Thus, the flexible loop seems to be an important distinguishing feature of the RIO kinase domain.

3.5. Autophosphorylation activity of the RIO kinases

The RIO proteins from yeast and archaea have been shown to undergo autophosphorylation in vitro [42–44]. At present, autophosphorylation sites for the Rio1 and Rio2 kinases from archaea have been identified, but equivalent sites in the yeast (or other) enzymes are unknown. In the case of Rio1, the autophosphorylation site, identified by phosphopeptide mapping and sequencing, was determined to be Ser108, a residue located on the flexible loop directly adjacent to the start of αC, and containing the sequence DMRRISPKEK [42]. Mutation of this residue to Ala results in a loss of autophosphorylation, but the mutant is capable of phosphorylating other substrates, as well as an inactive mutant of Rio1, with activity similar to that of the wild-type kinase. Thus, it would seem that lack of autophosphorylation does not affect the phosphorylation activity of the enzyme. This is in contrast with the report that dephosphorylated yeast Rio1 is nearly inactive [43]. Ser108 of *A. fulgidus* Rio1 is not conserved among the eukaryotic versions of the protein, so the autophosphorylation site(s) of these enzymes remain to be determined. For Rio2, the site of autophosphorylation was determined by phosphopeptide mapping and sequencing to be Ser128 [44], which is located also within the flexible loop, but this time near the end that is connected to β3, and within the sequence KVGHT**S**FKKVK. This serine is conserved among the eukaryotic Rio2 homologs and may represent a conserved regulatory site, but mutants have yet to be constructed and properly tested to confirm this hypothesis.

3.6. The hinge region of the RIO kinases

The flexible connection between the N-lobe and the C-lobe of the kinase domain around which they move in response to nucleotide binding has been called the hinge region. The Rio1 and Rio2 proteins exhibit a significant difference in this portion of the molecule which may translate into differences in how the two proteins bind ATP. The hinge region of canonical ePKs is typically 5- or 6-residue long and consist simply of an extended chain linker between the two lobes. A similar extended chain is also seen in Rio2. In Rio1, however, the linker region contains an insertion of five amino acids which allow for the formation

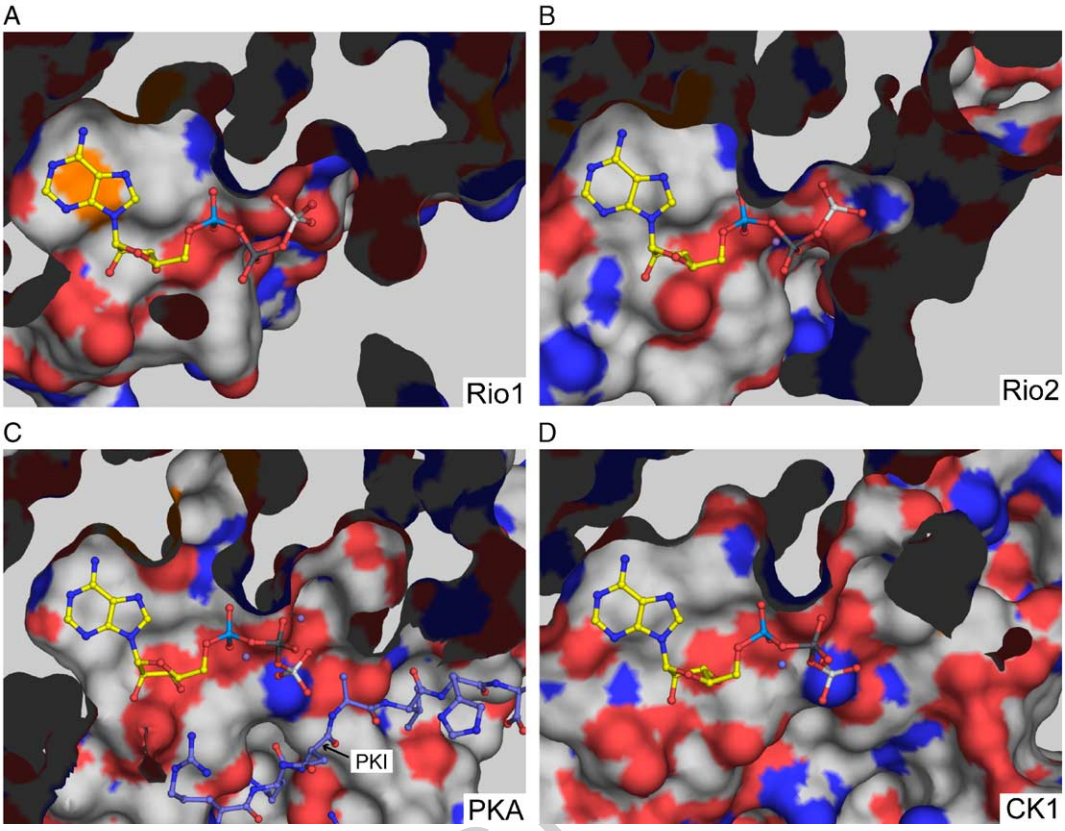


Fig. 4. The active site cavities of the RIO kinases. The active site cavities of the (A) Rio1, (B) Rio2, (C) PKA, and (D) CK1 are shown in surface representation, looking down from the top of the molecule. ATP and metal ion are shown in all four cases, and PKI is shown bound to PKA in (C) to indicate where a substrate would bind. The surface is colored by atom type (C—white, O—red, N—blue, S—orange).

of a β-hairpin connected by 3 hydrogen bonds (Fig. 2C). This results in closing off of the ATP-binding cavity and may result in a difference in affinity for ATP relative to Rio2. No equivalent β-hairpin has been seen in any other protein kinase known to date.

3.7. Additional domains of RIO kinases

Unexpectedly, the N-terminal Rio2-specific domain was found to contain a winged helix–loop–helix fold [41]. This fold type is seen primarily in DNA-binding proteins, but has been reported to mediate protein–protein and protein–RNA interactions as well [45–47]. Based on these findings, and given the role of Rio2 in rRNA processing in yeast, we have investigated the ability of Rio2 to bind nucleic acids. Analysis of the electrostatic properties of the surface of Rio2 shows that charge distribution is consistent with known nucleic acid-binding proteins, and fluorescence anisotropy experiments using labeled oligonucleotides indicate that Rio2 is indeed capable of binding single-stranded nucleic acids (LaRonde-LeBlanc, unpublished data). However, the domain does not contain many solvent-exposed residues that are conserved between the archaeal Rio2 and the eukaryotic Rio2 kinases, which would argue against a conserved nucleic acid binding site. The target of the binding may be sufficiently different between organisms to allow for this difference in the putative recognition residues. Indeed, if Rio2 proteins from metazoan organisms are aligned, much more conservation of surface residues is observed. For alignment of the human, mouse, rat, dog, frog, zebrafish, chicken, fly, and worm sequences, 58% of the residues are highly conserved or identical in the N-terminal winged helix domain of the Rio2 protein. If only the mammalian sequences are included, the sequence conservation is 81%. This includes the conservation of several basic amino acids in helix α3 and the wing of the winged helix domain which could potentially interact with the major groove of double-stranded nucleic acid. More data are required in order to determine the function of the winged helix domain, but the structure of *A. fulgidus* Rio2 has provided clues about which residues to probe to answer this question.

Although sequence alignments did not indicate that any other conserved domain(s) are present in the Rio1 proteins, the structure of *A. fulgidus* Rio1 [42] has identified an α-helix N-terminal to the RIO domain that appears to be conserved in other sequences as well. This helix is part of a conserved region that is only 14-residues long. The function of this helix cannot be elucidated by this structure, but the position of the helix relative to the RIO domain shifts if the Apo-Rio1 and the ATP-Mn-Rio1 are compared. An additional helix is also seen N-terminal to the conserved helix in the ATP complex, but not in the Apo structure. This helix does not appear to be conserved, and the significance of the conformational change is not apparent since this part of the molecule participates in different crystal contacts in the Apo and ATP-bound forms.

3.8. Two subfamilies of Rio1-like kinases

Only two RIO proteins, Rio1 and Rio2, have been identified in the *Saccharomyces cerevisiae* [35,43] and in *A. fulgidus* [1]. However, two genes encoding different Rio1-like enzymes can be distinguished in the genomes of mammals and other higher organisms. Their products have been named Rio1 and Rio3 [1], although the latter subfamily has not been so identified elsewhere. SudD, a product of a gene first identified in *Aspergillus nidulans* [16,48], is considered to be the defining member of the Rio3 subfamily. Through a comparison of structure-based sequence alignment, we noticed that the association between Rio3 and SudD may be, however, incorrect, since the latter enzyme appears to be more similar to mammalian Rio1 than to the putative Rio3. In particular, the Rio3 enzymes contain a unique and highly conserved N-terminal domain consisting of over 200 amino acids that is predicted to be highly helical. Thus far, no sequence homology to any known domains has been detected, but it is very highly conserved from human to flies. The presence of this divergent RIO kinase in higher eukaryotes suggests an additional function for the RIO kinases in these organisms, but the nature of such a function is not presently known. No similar domain is present in either SudD or in Rio1 from mammals or yeast, and thus SudD might be more similar to the Rio1 than to Rio3 enzymes. We have recently subcloned and expressed human Rio3 kinase (unpublished) and the structure of this enzyme may help in resolving the evolutionary relationships within the RIO family.

4. Conclusions

The studies of the RIO kinases have revealed the important structural characteristics that distinguish this group of serine kinases from their "typical" counterparts. The minimal RIO domain was revealed, as well as the structural features which distinguish Rio1 from Rio2 enzymes and provide the basis for distinct function. It is clear that the ATP binding pocket and the mode of substrate binding will be distinct for RIO kinases. Given that there is only one copy of each RIO subfamily member per organism, this should allow design of inhibitors with extreme specificity which would target a distinct pathway. For the RIO kinases, the target is indeed attractive, since ribosome biogenesis is an important requirement for tumor progression [49]. Production of massive amounts of daughter cells requires synthesis of large quantity of ribosomes. Shutting down a single RIO kinase will likely have the effect of stopping ribosome production.

There are still many unanswered questions regarding the structure of the RIO kinases. The structures that have been solved to date were obtained from the RIO kinases of an archaeal organism, which are perhaps the most divergent of all. As such, some of the structural features observed may not translate into the eukaryotic homologs. However, based on sequence analysis, we can conclude that certain structural features will indeed be present in eukaryotic RIO kinases as well. The lack of the activation loop, a surprising discovery because of its established role in peptide substrate binding and selectivity, is expected to also be a feature of the eukaryotic RIO kinases since their sequences contain no insertion between subdomains VII and IX when compared to the archaeal counterparts. A β-hairpin in the hinge region of the Rio1 kinases is also expected to be present in the eukaryotic Rio1 kinases, based on sequence comparisons. The P-loop, catalytic loop, and metal binding loop sequences are all very highly conserved between the archaeal and eukaryotic RIO kinases, so the ATP binding features described here may be the same, or very similar. The flexible loop and some of the sequence is conserved, and thus these may function in a similar fashion. However, the presence of a non-conserved autophosphorylation site in archaeal Rio1 may point to some differences in the way the flexible loop functions in the eukaryotic proteins. The eukaryotic RIO proteins also have the distinction of including a long stretch of conserved sequence beyond the C-terminus of the archaeal RIO kinases. This region has been identified as another domain and is termed "K-rich", due to a high percentage of lysine residues. This region may in fact be the missing subdomains X and XI of the ePK fold, or may be another domain entirely. Structural studies of eukaryotic RIO kinases are required in order to answer these questions.

How do peptide substrates bind to RIO kinases? With no activation loop present, there is no real way of obtaining this information by comparison with the ePK structures solved with bound peptides. It becomes even more challenging given the observation that the catalytic aspartate residues of the RIO kinases are more than 5.5 Å away from the γ-phosphate, compared to 3.8 Å in PKA. In addition, in the case of Rio2, the γ-phosphate is completely enclosed by an ordered part of the flexible loop and conformational changes would be required to allow access for phosphoryl transfer. Based on the position of conserved surface residues over a large area surrounding the active site of Rio2 and the lack of the activation loop, we initially proposed that RIO kinases may recognize a surface of a substrate protein rather than a peptide, in order to catalyze phosphorylation. However, the autophosphorylation sites determined for *A. fulgidus* Rio1 and Rio2 were both located within the flexible loop regions of the RIO domain, which suggests that the RIO kinases may indeed recognize an extended peptide. Further experiments are required to determine if peptides will indeed be accepted by the RIO kinases as substrates.

Although only two subfamilies of RIO kinases are generally recognized, we propose that the Rio3 kinases are sufficiently distinct from the Rio1 kinases to identify a third subfamily. The conserved N-terminal domain has no sequence homologs in any known proteins and thus its function is completely unknown. In addition, the bacterial RIO kinases may represent a fourth subfamily which may correspond to the progenitor of both the Rio1 and Rio2 kinases. Since the RIO kinases are essential proteins that have homologs in many prokaryotic and all eukaryotic organisms, including, for example, pathogens such as *Yersinia pestis*, this kinase may also become a tractable target for some pathogen-driven diseases. What remains to be deciphered is the exact function of each subfamily of RIO kinases in the organisms in which they are represented.

Acknowledgements

We would like to express our gratitude to Dr. Zbyszek Dauter for help in collecting and analyzing the crystallographic data for RIO kinases and Drs. Wolfhard Bandlow and Chris Vos for helpful discussions.

References

[1] G. Manning, D.B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, The protein kinase complement of the human genome, Science 298 (2002) 1912–1934.
[2] D.R. Knighton, J.H. Zheng, L.F. Ten Eyck, V.A. Ashford, N.H. Xuong, S.S. Taylor, J.M. Sowadski, Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase, Science 253 (1991) 407–414.
[3] D.R. Knighton, J.H. Zheng, L.F. Ten Eyck, N.H. Xuong, S.S. Taylor, J.M. Sowadski, Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase, Science 253 (1991) 414–420.
[4] S.K. Hanks, T. Hunter, Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification, FASEB J. 9 (1995) 576–596.
[5] L.N. Johnson, E.D. Lowe, M.E.M. Noble, D.J. Owen, The structural basis for substrate recognition and control by protein kinases, FEBS Lett. 430 (1998) 1–11.
[6] L.N. Johnson, M.E. Noble, D.J. Owen, Active and inactive protein kinases: structural basis for regulation, Cell 85 (1996) 149–158.
[7] D. Drennan, A.G. Ryazanov, Alpha-kinases: analysis of the family and comparison with conventional protein kinases, Prog. Biophys. Mol. Biol. 85 (2004) 1–32.
[8] A.G. Ryazanov, M.D. Ward, C.E. Mendola, K.S. Pavur, M.V. Dorovkov, M. Wiedmann, H. Erdjument-Bromage, P. Tempst, T.G. Parmer, C.R. Prostko, F.J. Germino, W.N. Hait, Identification of a new class of protein kinases represented by eukaryotic elongation factor-2 kinase, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 4884–4889.
[9] G.P. Cote, X. Luo, M.B. Murphy, T.T. Egelhoff, Mapping of the novel protein kinase catalytic domain of Dictyostelium myosin II heavy chain kinase A, J. Biol. Chem. 272 (1997) 6846–6849.
[10] L.W. Runnels, L. Yue, D.E. Clapham, TRP-PLIK, a bifunctional protein with kinase and ion channel activities, Science 291 (2001) 1043–1047.
[11] H. Yamaguchi, M. Matsushita, A.C. Nairn, J. Kuriyan, Crystal structure of the atypical protein kinase domain of a TRP channel with phosphotransferase activity, Mol. Cell 7 (2001) 1047–1057.
[12] J.F. Beeler, W.J. LaRochelle, M. Chedid, S.R. Tronick, S.A. Aaronson, Prokaryotic expression cloning of a novel human tyrosine kinase, Mol. Cell. Biol. 14 (1994) 982–988.
[13] A. Rohwer, W. Kittstein, F. Marks, M. Gschwendt, Cloning, expression and characterization of an A6-related protein, Eur. J. Biochem. 263 (1999) 518–525.
[14] R.T. Abraham, PI 3-kinase related kinases: 'big' players in stress-induced signaling pathways, DNA Repair 3 (2004) 883–887.
[15] E.H. Walker, O. Perisic, C. Ried, L. Stephens, R.L. Williams, Structural insights into phosphoinositide 3-kinase catalysis and signalling, Nature 402 (1999) 313–320.
[16] C.J. Leonard, L. Aravind, E.V. Koonin, Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily, Genome Res. 8 (1998) 1038–1047.
[17] G.V. Denis, M.R. Green, A novel, mitogen-activated nuclear kinase is related to a *Drosophila* developmental regulator, Genes Dev. 10 (1996) 261–271.
[18] B. Florence, D.V. Faller, You bet-cha: a novel family of transcriptional regulators, Front Biosci. 6 (2001) D1008–D1018.
[19] R.J. Greenwald, J.R. Tumang, A. Sinha, N. Currier, R.D. Cardiff, T.L. Rothstein, D.V. Faller, G.V. Denis, E mu-BRD2 transgenic mice develop B-cell lymphoma and leukemia, Blood 103 (2004) 1475–1484.
[20] G.V. Denis, C. Vaziri, N. Guo, D.V. Faller, RING3 kinase transactivates promoters of cell cycle regulatory genes through E2F, Cell Growth Differ. 11 (2000) 417–424.
[21] S. Faderl, M. Talpaz, Z. Estrov, S. O'Brien, R. Kurzrock, H.M. Kantarjian, The biology of chronic myeloid leukemia, N. Engl. J. Med. 341 (1999) 164–172.
[22] I.K. Hariharan, J.M. Adams, cDNA sequence for human bcr, the gene that translocates to the abl oncogene in chronic myeloid leukaemia, EMBO J. 6 (1987) 115–119.
[23] Y. Maru, O.N. Witte, The BCR gene encodes a novel serine/threonine kinase activity within a single exon, Cell 67 (1991) 459–468.
[24] G. Radziwill, R.A. Erdmann, U. Margelisch, K. Moelling, The Bcr kinase downregulates Ras signaling by phosphorylating AF-6 and binding to its PDZ domain, Mol. Cell. Biol. 23 (2003) 4663–4672.
[25] C.C. Smith, Y.X. Yu, M. Kulka, L. Aurelian, A novel human gene similar to the protein kinase (PK) coding domain of the large subunit of herpes simplex virus type 2 ribonucleotide reductase (ICP10) codes for a serine-threonine PK and is expressed in melanoma cells, J. Biol. Chem. 275 (2000) 25690–25699.
[26] J.W. Nelson, J. Zhu, C.C. Smith, M. Kulka, L. Aurelian, ATP and SH3 binding sites in the protein kinase of the large subunit of herpes simplex virus type 2 of ribonucleotide reductase (ICP10), J. Biol. Chem. 271 (1996) 17021–17027.
[27] G. Kappe, P. Verschuure, R.L. Philipsen, A.A. Staalduinen, B.P. Van de, W.C. Boelens, W.W. De Jong, Characterization of two novel human small heat shock proteins: protein kinase-related HspB8 and testis-specific HspB9, Biochim. Biophys. Acta 1520 (2001) 1–6.
[28] Q. Tian, J. Taupin, S. Elledge, M. Robertson, P. Anderson, Fas-activated serine/threonine kinase (FAST) phosphorylates TIA-1 during Fas-mediated apoptosis, J. Exp. Med. 182 (1995) 865–874.
[29] L.R. Stepp, F.H. Pettit, S.J. Yeaman, L.J. Reed, Purification and properties of pyruvate dehydrogenase kinase from bovine kidney, J. Biol. Chem. 258 (1983) 9454–9458.
[30] M.S. Patel, L.G. Korotchkina, Regulation of mammalian pyruvate dehydrogenase complex by phosphorylation: complexity of multiple phosphorylation sites and kinases, Exp. Mol. Med. 33 (2001) 191–197.
[31] C.N. Steussy, K.M. Popov, M.M. Bowker-Kinley, R.B. Sloan Jr., R.A. Harris, J.A. Hamilton, Structure of pyruvate dehydrogenase kinase. Novel folding pattern for a serine protein kinase, J. Biol. Chem. 276 (2001) 37443–37450.
[32] R. Dikstein, S. Ruppert, R. Tjian, TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74, Cell 84 (1996) 781–790.
[33] T. O'Brien, R. Tjian, Functional analysis of the human TAFII250 N-terminal kinase domain, Mol. Cell 1 (1998) 905–911.
[34] R.A. Fraser, D.J. Heard, S. Adam, A.C. Lavigne, B. Le Douarin, L. Tora, R. Losson, C. Rochette-Egly, P. Chambon, The putative cofactor TIF1alpha is a protein kinase that is hyperphosphorylated upon interaction with liganded nuclear receptors, J. Biol. Chem. 273 (1998) 16199–16204.
[35] M. Angermayr, A. Roidl, W. Bandlow, Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression, Mol. Microbiol. 44 (2002) 309–324.
[36] A. Krupa, N. Srinivasan, Lipopolysaccharide phosphorylating enzymes encoded in the genomes of Gram-negative bacteria are related to the eukaryotic protein kinases, Protein Sci. 11 (2002) 1580–1584.
[37] T.H. Geerlings, A.W. Faber, M.D. Bister, J.C. Vos, H.A. Raue, Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in *Saccharomyces cerevisiae*, J. Biol. Chem. 278 (2003) 22537–22545.
[38] E. Vanrobays, P.E. Gleizes, C. Bousquet-Antonelli, J. Noaillac-Depeyre, M. Caizergues-Ferrer, J.P. Gelugne, Processing of 20 S pre-rRNA to 18 S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein, EMBO J. 20 (2001) 4204–4213.
[39] E. Vanrobays, J.P. Gelugne, P.E. Gleizes, M. Caizergues-Ferrer, Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in *Saccharomyces cerevisiae*, Mol. Cell. Biol. 23 (2003) 2083–2095.
[40] G. Giaever, A.M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A.P. Arkin, A. Astromoff, M. El

Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K.D. Entian, P. Flaherty, F. Foury, D.J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J.H. Hegemann, S. Hempel, Z. Herman, D.F. Jaramillo, D.E. Kelly, S.L. Kelly, P. Kotter, D. LaBonte, D.C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S.L. Ooi, J.L. Revuelta, C.J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D.D. Shoemaker, S. Sookhai-Mahadeo, R.K. Storms, J.N. Strathern, G. Valle, M. Voet, G. Volckaert, C.Y. Wang, T.R. Ward, J. Wilhelmy, E.A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J.D. Boeke, M. Snyder, P. Philippsen, R.W. Davis, M. Johnston, Functional profiling of the *Saccharomyces cerevisiae* genome, Nature 418 (2002) 387–391.

[41] N. LaRonde-LeBlanc, A. Wlodawer, Crystal structure of *A. fulgidus* Rio2 defines a new family of serine protein kinases, Structure 12 (2004) 1585–1594.

[42] N. LaRonde-LeBlanc, T. Guszczynski, T. Copeland, A. Wlodawer, Structure and activity of the atypical serine kinase Rio1. FEBS J. (in press).

[43] M. Angermayr, W. Bandlow, RIO1, an extraordinary novel protein kinase, FEBS Lett. 524 (2002) 31–36.

[44] N. LaRonde-LeBlanc, T. Guszczynski, T.D. Copeland, A. Wlodawer, Autophosphorylation of A. fulgidus Rio2 and crystal structures of its nucleotide–metal ion complexes, FEBS J. (in press).

[45] K.S. Gajiwala, S.K. Burley, Winged helix proteins, Curr. Opin. Struct. Biol. 10 (2000) 110–116.

[46] C. Wolberger, R. Campbell, New perch for the winged helix, Nat. Struct. Biol. 7 (2000) 261–262.

[47] G. Dong, G. Chakshusmathi, S.L. Wolin, K.M. Reinisch, Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch, EMBO J. 23 (2004) 1000–1007.

[48] P. Anaya, S.C. Evans, C. Dai, G. Lozano, G.S. May, Isolation of the *Aspergillus nidulans* sudD gene and its human homologue, Gene 211 (1998) 323–329.

[49] D. Ruggero, P.P. Pandolfi, Does the ribosome translate cancer? Nat. Rev., Cancer 3 (2003) 179–192.