The PB1 domain and the PC motif-containing region are structurally similar protein binding modules

Sosuke Yoshinaga¹, Motoyuki Kohjima², Kenji Ogura¹, Masashi Yokochi^{1,3}, Ryu Takeya², Takashi Ito^{4,5}, Hideki Sumimoto^{2,3} and Fuyuhiko Inagaki^{1,3,6}

¹Department of Structural Biology, Graduate School of Pharmaceutical Sciences, Hokkaido University, N12, W6, Kita-ku, Sapporo 060-0812,

²Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, ³CREST, Japan Science and Technology, ⁴Division of Genome Biology, Cancer Research Institute, Kanazawa University, 13-1 Takaramachi, Kanazawa, Ishikawa 920-0934 and ⁵Department of Computational Biology, Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan

⁶Corresponding author
e-mail: finagaki@pharm.hokudai.ac.jp

The PC motif is evolutionarily conserved together with the PB1 domain, a binding partner of the PC motif-containing protein. For interaction with the PB1 domain, the PC motif-containing region (PCCR) comprising the PC motif and its flanking regions is required. Because the PB1 domain and the PCCR are novel binding modules found in a variety of signaling proteins, their structural and functional characterization is crucial. Bem1p and Cdc24p interact through the PB1–PCCR interaction and regulate cell polarization in budding yeast. Here, we determined a tertiary structure of the PCCR of Cdc24p by NMR. The tertiary structure of the PCCR is similar to that of the PB1 domain of Bem1p, which is classified into a ubiquitin fold. The PC motif portion takes a compact $\beta\beta\alpha$ -fold, presented on the ubiquitin scaffold. Mutational studies indicate that the PB1–PCCR interaction is mainly electrostatic. Based on the structural information, we group the PB1 domains and the PCCRs into a novel family, named the PB1 family. Thus, the PB1 family proteins form a specific dimer with each other.

Keywords: Bem1p/Cdc24p/cell polarity/NADPH oxidase/NMR structure

Introduction

The PC motif consists of 28 highly conserved acidic and hydrophobic amino acid residues (Nakamura *et al.*, 1998). The name is derived from the fact that the motif occurs in mammalian p40^{phox} (a cytosolic factor of the NADPH oxidase in phagocytes) (Wientjes *et al.*, 1993) and budding yeast Cdc24p (a guanine nucleotide-exchange factor for a small GTPase Cdc42p) (Miyamoto *et al.*, 1987). The PC motif is also known as OPR (octicosapeptide repeat) (Ponting, 1996) or AID (atypical PKC-interaction domain) (Moscat and Diaz-Meco, 2000), which are collectively

named the OPCA (OPR, PC and AID) motif (Ponting *et al.*, 2002). A variety of cellular signaling proteins containing the PC motif have been reported so far (Figure 1A): Scd1 (a fission yeast ortholog of Cdc24p) (Chang *et al.*, 1994), aPKC (a Ser/Thr kinase involved in cell proliferation, differentiation and survival) (Ono *et al.*, 1989; Selbie *et al.*, 1993; Akimoto *et al.*, 1994), PAR-6 (a scaffold protein involved in cell polarization) (Watts *et al.*, 1996; Hung and Kemphues, 1999; Joberty *et al.*, 2000; Noda *et al.*, 2001), p62 (a scaffold protein involved in cell survival) (Devergne *et al.*, 1996; Joung *et al.*, 1996; Marcus *et al.*, 1996), MEK5 (a MAPKK involved in EGF-induced cell proliferation and T-cell activation) (English *et al.*, 1995; Zhou *et al.*, 1995), TFG (a protein fused to a tyrosine kinase domain of TrkA in papillary thyroid carcinomas or that of ALK in anaplastic large-cell lymphomas) (Greco *et al.*, 1995; Mencinger *et al.*, 1997; Hernández *et al.*, 1999, 2002) and NBR1 (a protein involved in neural development) (Campbell *et al.*, 1994; Whitehouse *et al.*, 2002). However, the functional roles of the PC motif have remained elusive.

Analysis of the binding proteins of the PC motif-containing region (PCCR) of p40^{phox} and Cdc24p has led to the identification of a novel protein binding domain, the PB1 domain (Ito *et al.*, 2001), which is named after p67^{phox} (Leto *et al.*, 1990) and Bem1p (Chenevert *et al.*, 1992), binding partners of p40^{phox} and Cdc24p, respectively. Thus, the PB1 domain and the PCCR are identified as novel protein binding modules involved in a variety of cellular signaling proteins of origins ranging from yeast to human.

The biological implications of the PB1–PCCR interaction have been studied in several systems. In budding yeast cells, Bem1p mutant proteins, which are defective in the PB1–PCCR interaction, cannot maintain Cdc24p at the site of polarized growth. As a consequence, such cells lack apical bud growth and the formation of a mating projection (Ito *et al.*, 2001; Butty *et al.*, 2002), while in stimulated K562 cells the p67^{phox} and p40^{phox} inactive mutant proteins reduce the membrane translocation efficiency of a cytosolic regulatory complex in the phagocyte NADPH oxidase so that superoxide production, essential for the disinfection of microorganisms, is largely reduced (Kuribayashi *et al.*, 2002). Recently, an interaction between aPKC and PAR-6, proteins that contain the PB1 domain or the PCCR (Noda *et al.*, 2003), has been identified as crucial for the control of cell polarization in many cell types: epithelial cells, hippocampal neurons, astrocytes, *Drosophila melanogaster* neuroblasts and *Caenorhabditis elegans* embryos (Ohno, 2001; Etienne-Manneville and Hall, 2003). Thus, the PB1–PCCR interaction has an important role in the formation of protein complexes required for cell polarization and their recruitment to the functional sites in a variety of cells.

Fig. 1. (A) Sequence alignment of the PC motifs: #, hydrophobic amino acid; %, hydrophilic amino acid; s.c., *Saccharomyces cerevisiae* (budding yeast); s.p., *Schizosaccharomyces pombe* (fission yeast); h.s., *Homo sapiens* (human). (B) Domain structure and part of the amino acid sequence of Cdc24p. Cdc24p includes a calponin homology domain (CH), a Dbl homology domain (DH), a pleckstrin homology domain (PH), a Ser/Thr-rich region (ST rich) and the PCCR from the N-terminus. The N-terminal side of the PCCR includes hydrophobic residues shaded in yellow.

The structural characterization of the PB1 domain and the PCCR is helpful for understanding the strict control of cellular signaling. In a previous paper, we reported the NMR structure of the PB1 domain of Bem1p (abbreviated to Bem1 PB1). It presents a conserved lysine residue essential for the interaction with the PCCR of Cdc24p (abbreviated to Cdc24 PCCR) on a ubiquitin fold (Terasawa *et al.*, 2001). Further, we reported the secondary structure and the topology of Cdc24 PCCR (780–854), a minimum sequence for the interaction with Bem1 PB1 on the basis of nuclear Overhauser effect (NOE) connectivities observed among the main chain proton resonances and CSI analysis (Terasawa *et al.*, 2001).

We initially determined the NMR structure of Cdc24 PCCR (780–854) (unpublished data). The overall structure is most similar to Bem1 PB1, the binding partner of Cdc24 PCCR, but Cdc24 PCCR (780–854) is lacking in the first β -strand characteristic of a ubiquitin fold. This raised the concern about whether the structure was an artifact due to truncation of the N-terminal residues. We searched for the sequence in the Ser/Thr-rich (ST-rich) region preceding Cdc24 PCCR (780–854) and identified the domain boundary based on the hydrophobicity of the sequence (Figure 1B). Subsequently, we determined the

topology of this N-terminally extended Cdc24 PCCR including the residues 761–854, designated as Cdc24 PCCR (761–854). Interestingly, the N-terminally extended region is wedged into the β -sheet of Cdc24 PCCR (780–854) and forms the first β -strand so that the topology of Cdc24 PCCR (761–854) is similar to that of Bem1 PB1 or proteins with a ubiquitin fold. This observation, together with the sequence alignment of the PB1 domains and the PCCRs, resulted in the grouping of these domains into the PB1 family (Ponting *et al.*, 2002).

Here, we report the NMR structure of Cdc24 PCCR (761–854). The unique PC motif portion of Cdc24 PCCR forms a compact $\beta\beta\alpha$ -fold, which is presented on the scaffold of the ubiquitin fold. Mutational studies indicate that the PB1–PCCR interaction requires mutually different protein surfaces, where the PC motif is used as a binding site for the PB1 domain. Thus, the PB1 family proteins form an asymmetric dimer with each other.

Results and discussion

Structure of Cdc24 PCCR (761–854)

The assignments of the NMR resonances derived from the ^1H , ^{15}N and ^{13}C nuclei in Cdc24 PCCR (761–854) were

Table I. Experimental restraints and structural statistics for the 20 lowest energy structures

Number of experimental restraints	
Total unambiguous distance restraints	1707
Intraresidual	686
Sequential ($li - jl = 1$)	238
Medium range ($2 \leq li - jl \leq 4$)	156
Long range ($5 \leq li - jl$)	627
Ambiguous distance restraints	394
ϕ angle restraints	44
Hydrogen bonds	35
Structural statistics	
R.m.s.d. from mean coordinate ^a	
Backbone (N, C α , C')	0.35 Å
Heavy atoms	0.79 Å
R.m.s.d. from experimental restraints	
NOE distances	0.003 $\pm$ 0.000 Å
Dihedral angles	0.012 $\pm$ 0.015°
R.m.s.d. from idealized geometry	
Bonds	0.001 $\pm$ 0.000 Å
Angles	0.234 $\pm$ 0.001°
Impropers	1.00 $\pm$ 0.04°
Final energies	
E_{total}	50.18 $\pm$ 0.60 kcal/mol
E_{bonds}	0.65 $\pm$ 0.04 kcal/mol
E_{angles}	23.52 $\pm$ 0.17 kcal/mol
E_{improper}	1.24 $\pm$ 0.09 kcal/mol
E_{vdw}	23.90 $\pm$ 0.71 kcal/mol
E_{NOE}	0.87 $\pm$ 0.23 kcal/mol
E_{cdih}	0.00 $\pm$ 0.00 kcal/mol
Ramachandran analysis ^{a,b}	
Residues in most favored regions	86.8%
Residues in additionally allowed regions	13.2%
Residues in generously allowed regions	0.0%
Residues in disallowed regions	0.0%

^aCalculated with the residues except the protease cleavage site, flexible loops (769–778 and 803–815) and C-terminal (854) regions.

^bNon-Gly/Pro residues for the mean coordinate. The program PROCHECK-NMR (Laskowski *et al.*, 1996) was used to assess the quality of the structure.

made with a suite of triple-resonance multi-dimensional NMR spectra (Cavanagh *et al.*, 1996). The structure of Cdc24 PCCR (761–854) was determined on the basis of 2101 NOE-derived interproton distance restraints, 44 ϕ dihedral angle restraints and 35 hydrogen bond restraints (Table I) that were obtained as described in Materials and methods. A total of 100 structures were calculated, and the 20 lowest-energy structures were selected. Residues 769–778 and 803–815 are disordered due to a lack of long-range NOEs. As for the residues 761–768, 779–802 and 816–853, the root mean square deviation (r.m.s.d.) from the mean coordinate is 0.35 Å for the backbone heavy atoms, indicating that the backbone of Cdc24 PCCR (761–854) is well defined, as illustrated in Figure 2A. Cdc24 PCCR (761–854) contains two α -helices (α 1, residues 791–802; α 2, 831–842) and a mixed β -sheet that consists of five strands (β 1, residues 762–768; β 2, 779–785; β 3, 816–819; β 4, 825–828; β 5, 848–853) (Figure 2A–C). The inner two strands, β 1 and β 5, are parallel, while β 2, β 3 and β 4 run in an antiparallel manner (Figure 2B and C). The N-terminal extension (shown in green in Figure 2C) forms β 1, a flexible loop and the

N-terminal part of β 2. Thus, Cdc24 PCCR (761–854) has a canonical ubiquitin fold.

The RMSDs between Cdc24 PCCR (761–854) and Bem1 PB1 (Terasawa *et al.*, 2001) or ubiquitin (Cornilescu *et al.*, 1998) are 2.2 Å or 2.5 Å for the backbone atoms of 66 structurally equivalent residues, even though the overall sequence similarity is below the level of significance (Chothia and Lesk, 1986). The region corresponding to the PC motif is the most variable region in the ubiquitin fold. Bem1 PB1 has a flexible loop and is lacking in one β -strand in this region, while ubiquitin has a smaller number of residues than Cdc24 PCCR (761–854) and a 3_{10} -helix instead of an α -helix (Figure 3). Therefore, the PC motif is characteristic of PCCR in terms of the structural aspects.

The unique PC motif portion of Cdc24p (Ile813–Met840) consists of β 3, β 4 and α 2, and forms a compact three-dimensional (3D) structure. Each of the side chains of Tyr818 and Phe825 is located on the mutually opposite side of the β -sheet and forms part of the structural core (enclosed by red rectangles in Figure 2B). The tyrosine and phenylalanine residues in the PC motif are conserved and type-conserved, respectively (Figure 1). The five consecutive residues Asp820–Glu821–Asp822–Gly823–Asp824, corresponding to a linker between β 3 and β 4, form a unique turn called a type I + G1 β -bulge (Sibanda and Thornton, 1985; Milner-White, 1987; Sibanda *et al.*, 1989; Chan *et al.*, 1993). Each of the amide groups of Asp820 and Gly823 forms a hydrogen bond with the carbonyl group of Asp824 and Asp820. Because of the steric hindrance, this type of turn prefers a glycine residue in the fourth position, which is conserved in the PC motifs (Figure 1). The characteristic arrangement of the residues on the turn partly distorts the β -sheet so that the acidic side chains of Asp820, Glu821, Asp822 and Asp824 are presented on the same side and form a highly acidic surface (Figure 2B). These acidic residues are conserved in the PC motifs (Figure 1).

Binding mode between Bem1 PB1 and Cdc24 PCCR

In our previous study, we reported the chemical shift differences of the backbone amide ^1H and ^{15}N resonances between the free and bound states of Cdc24 PCCR (780–854) together with the binding studies of its single mutants substituted to an alanine residue toward Bem1 PB1 by an *in vitro* pull-down assay (Terasawa *et al.*, 2001). Both studies indicated that the region containing the PC motif of Cdc24 PCCR (780–854) is the direct interaction site for Bem1 PB1. Here, we determined the tertiary structures of Cdc24 PCCR (780–854) (unpublished data) and Cdc24 PCCR (761–854) by NMR. Interestingly, the molecular surfaces formed by the PC motif of Cdc24 PCCR (780–854) and (761–854) are quite similar, regardless of the truncation of the first β -strand in Cdc24 PCCR (780–854). The surface contains two clusters of conserved acidic residues separated from each other in both sequence and 3D structure (Figures 1A and 4C, right). One cluster includes Asp820, Glu821, Asp822 and Asp824 (designated Cluster A1) located on the turn, the type I + G1 β -bulge, and the other cluster includes Asp831, Glu832 and Asp833 (designated Cluster A2) which form the N-terminal part of α 2. A hydrophobic region is located between the two acidic clusters. Therefore, both

Fig. 2. (A) A stereo pair of overlays of the 20 lowest-energy structures of Cdc24 PCCR (761–854) on the average coordinate position for the backbone heavy atoms of the residues 761–768, 779–802 and 816–853. (B) Ribbon diagrams of the lowest-energy structure of Cdc24 PCCR (761–854). The region corresponding to the PC motif is surrounded by red rectangles. The side chains of Y818 and D820–F825 are shown as ball-and-stick models. The structure on the right is related to the structure on the left by 270° rotation around the vertical axis. (C) Topology diagram of Cdc24 PCCR (761–854). The N-terminal extension shown in green forms $\beta 1$, a flexible loop and a part of $\beta 2$ in Cdc24 PCCR (761–854).

Cdc24 PCCR (780–854) and (761–854) are considered to interact with Bem1 PB1 through this molecular surface, which is consistent with the experimental evidence that Cdc24 PCCR (780–854) is a minimum sequence required

for binding to Bem1 PB1 (Ito *et al.*, 2001) and both proteins interact with Bem1 PB1 with similar affinity.

Subsequently, we performed additional mutational experiments on the surface residues to elucidate the

Fig. 3. Comparison of the structures between Cdc24 PCCR (761–854) and Bem1 PB1 or ubiquitin. The structure of the PC motif portion surrounded by a red rectangle is characteristic of Cdc24 PCCR. The PDB accession codes are 1IPG (Bem1 PB1) and 1D3Z (ubiquitin).

binding mode between Bem1 PB1 and Cdc24 PCCR (761–854). The binding affinities of single mutants of Cdc24 PCCR (761–854) toward Bem1 PB1 were measured by an *in vitro* pull-down assay (Figure 4B) as described in Materials and methods, where each of Asp833, Val836, Glu839 and Val826 was substituted to an alanine residue, while Tyr818 was substituted to a phenylalanine residue. These results are mapped on the molecular surface of Cdc24 PCCR (761–854) together with the previous alanine-scanning results (Terasawa *et al.*, 2001) (Figure 4C, right).

D820A and D824A mutants on Cluster A1 decrease relative binding affinities to less than 10% of that of the wild type and E821A decreases it to about 60% (Terasawa *et al.*, 2001). Thus, the acidic residues in Cluster A1 are responsible for the interaction with Bem1 PB1. As for Cluster A2, D833A decreases the relative binding affinity to less than 10% (Figure 4B), while D831A has a binding affinity comparable to that of the wild type (Terasawa *et al.*, 2001). Asp833 is presented on the same side of Cluster A1 residues, while Asp831 is on the skew. As for the hydrophobic residue mutants, V826A decreases the relative binding affinity to 10% of that of the wild type and F825A decreases it to 60%, while V836A and V827A only decrease it to 80% (Figure 4B) (Terasawa *et al.*, 2001). Thus, the hydrophobic region also contributes to the binding affinity with Bem1 PB1. As for the basic residue mutants, K815A, K817A and R852A show binding affinities similar to that of the wild type (Terasawa *et al.*, 2001) and thus the contributions of the basic residues might be small. The mutant Y818F decreases the relative binding affinity to 70% (Figure 4B), and thus the hydroxyl group of Tyr818 might form a hydrogen bond with Bem1 PB1.

Previously, we proposed the Cdc24 PCCR binding surface of Bem1 PB1 based on the chemical shift differences between the free and bound states of Bem1 PB1 (Terasawa *et al.*, 2001). The surface comprised of β 1, β 2, β 4 and a C-terminal region of α 1 contains two basic clusters in accordance with the two acidic clusters of

Cdc24 PCCR (Figure 4C, left). One cluster includes Lys480, Lys482, Lys521, Lys543 and Lys545 (designated Cluster B1), and the other cluster includes Lys506 and Arg510 (designated Cluster B2). The binding affinities of the single mutants of Bem1 PB1 toward Cdc24 PCCR (761–854) were measured by an *in vitro* pull-down assay (Figure 4A). The residues on the Cdc24 PCCR binding surface of Bem1 PB1 were substituted to an alanine residue, except for Ala491 which was substituted to an asparagine residue. These results are mapped on the molecular surface of Bem1 PB1 (Figure 4C, left).

The mutant K482A markedly decreases the relative binding affinity (Figure 4A), and the mutants K480A, I489A, A491N and R510A moderately decrease the relative binding affinities (Figure 4A). Then, the mutants Y484A, M493A, K506A, K521A, K543A and K545A show binding affinities comparable to that of the wild type (Figure 4A). Although many basic residues are presented on the surface, only three residues (Lys482, Lys480 and Arg510) are responsible for the binding to Cdc24 PCCR. Lys482 is conserved in the PB1 domain, and Lys480 and Arg510 are conserved in the Bem1p orthologs. The essential role of Lys480 and Lys482 was previously shown *in vivo*; the yeast cells bearing a Bem1p mutant (K482A) protein displayed temperature-sensitive growth and a bilateral mating defect (Ito *et al.*, 2001) and those bearing a Bem1p mutant (K480E) protein could not maintain Cdc24p at sites of polarized growth and showed defects in apical bud growth (Butty *et al.*, 2002). It should be noted that the hydrophobic residues Ile489 and Ala491 between the two basic clusters also make significant contributions to the binding affinity with Cdc24 PCCR.

All the mutational effects and the structural studies taken together indicate that the electrostatic distribution and the location of hydrophobic region of the binding surfaces of both proteins are complementary to each other when two proteins are overlaid so that Clusters A1 and A2 dock Clusters B1 and B2, respectively (Figure 4C). This suggests that the electrostatic and hydrophobic interactions are the driving force behind the binding of Bem1 PB1 and Cdc24 PCCR.

Classification of the PB1 domain and the PCCR

Now that we have determined the three dimensional structure of Cdc24 PCCR (761–854), we can align the amino acid sequences of the PB1 domains and the PCCRs based on the 3D structures of Bem1 PB1 and Cdc24 PCCR (Figure 5). Although the PB1 domains and the PCCRs have low sequence identity (10–20%), we can group both domains into the PB1 family (Ponting *et al.*, 2002). The conserved or type-conserved residues are located on the secondary structural elements. Most of these residues are supposed to be essential for the maintenance of the ubiquitin fold.

It should be noted that the PB1 family proteins use distinct binding sites on the ubiquitin fold for heterodimer formation as was exemplified in the interaction between Cdc24p and Bem1p. The Bem1 PB1 binding site of Cdc24 PB1 (enclosed in magenta in Figure 6A and B) comprises the turn between β 3 and β 4, β 4 and α 2 included in the PC motif. The Cdc24 PB1 binding site of Bem1 PB1 (enclosed in green in Figure 6A and B) consists of β 1, β 2 and the C-terminal region of α 1. Thus, the PB1 family

Fig. 4. (A) Mutational effects of Bem1 PB1 on complex formation with Cdc24 PCCR (761–854) examined by *in vitro* pull-down assays. MBP-Cdc24 PCCR (761–854) was incubated with GST-Bem1 PB1 mutants carrying the indicated amino acid substitution. Proteins pulled down with glutathione–Sepharose-4B beads were subjected to SDS–PAGE and stained with Coomassie brilliant blue. To estimate the amounts of protein on the gel, densitometric analysis was carried out using the software NIH Image. The values represent the mean $\pm$ SD of six independent experiments. (B) Mutational effects of Cdc24 PCCR (761–854) on complex formation with Bem1 PB1 examined by *in vitro* pull-down assays. The assays were performed as in (A). (C) Electrostatic surface potential of the putative binding sites of Bem1 PB1 and Cdc24 PCCR (761–854). Two basic clusters (B1 and B2) of Bem1 PB1 and two acidic clusters (A1 and A2) of Cdc24 PCCR (761–854) are enclosed by blue and red ovals, respectively. Mutational results shown in (A) and (B) together with the previous results on the alanine-scanning mutagenesis (Terasawa *et al.*, 2001) are mapped on the surface. The residues with markedly decreased affinity are in yellow, with moderately decreased affinity in green and no effect in cyan. The orientation of Cdc24 PCCR (761–854) is the same as in Figure 2B (right).

proteins are classified into two types according to the binding sites used for PB1 dimer formation: type I (Cdc24 PB1 type) and type II (Bem1 PB1 type) (Figures 5 and 6A and B). The residues on the PC motif of p40^{phox} PB1 are required for the interaction with p67^{phox} PB1 (Nakamura *et al.*, 1998), and thus p40^{phox} PB1 belongs to type I. Likewise, the conserved residue Lys355, located on β 1 of p67^{phox} PB1, is essential for the interaction with p40^{phox}

PB1 (Ito *et al.*, 2001) and, accordingly, p67^{phox} PB1 belongs to type II.

Interestingly, both aPKC PB1 and its binding partners, PAR-6 PB1, p62 PB1 and MEK5 PB1 contain the PC motif (Figure 1A and 5). Recently, interactions between these PB1 domains were investigated using a mutational approach. The PC motif residues of aPKC PB1 and the conserved lysine residue located on β 1 of PAR-6 PB1 or

Fig. 5. 3D-structure-based sequence alignment of the PB1 family proteins. The classification of type I and type II is based on the binding sites used for interaction with the cognate binding partner (Figure 6A and B): #, hydrophobic amino acid; %, hydrophilic amino acid; /, N- or C-terminus of the entire protein; s.c., *Saccharomyces cerevisiae* (budding yeast); s.p., *Schizosaccharomyces pombe* (fission yeast); h.s., *Homo sapiens* (human). The circles shown above the amino acid sequences denote the residues mutated for the binding studies (Terasawa *et al.*, 2001) (Figure 4A and B). The residues with markedly decreased affinity are in yellow, with moderately decreased affinity in green and no effect in cyan. The secondary structures shown above the amino acid sequences denote that of Cdc24p and Bem1p, respectively. p40^{phox} has the C-terminal extension (330) DNYRVYNTMP (339), following the above sequence. The amino acid sequence and the secondary structures of ubiquitin are shown for comparison.

p62 PB1 are required for the aPKC-PAR-6 or aPKC-p62 interactions (Gao *et al.*, 2002; Lamark *et al.*, 2003; Noda *et al.*, 2003), while the conserved lysine residues of aPKC PB1 or p62 PB1 are required for the MEK5-aPKC or MEK5-p62 interactions (Noda *et al.*, 2003). Thus, aPKC PB1 uses either of the two interaction sites in accordance with the binding partner. Moreover, homodimerization of p62 PB1, NBR1 PB1 or TFG PB1 was recently observed (Puls *et al.*, 1997; Lamark *et al.*, 2003; Noda *et al.*, 2003), suggesting the interaction between the PC motif residues and the conserved lysine residue. Altogether, some PB1 domains can be assigned to both type I and type II in accordance with the binding partner (Figure 5).

Several non-conserved residues presented on the binding surfaces of the PB1 domains are supposed to be specificity determinants for cognate pairs of the PB1 domains. As for the PB1 domain type I, the residues Glu821, Phe825, Val826 and Glu832 in Cdc24p are not conserved, and, as for the PB1 domain type II, the residues Lys480, Ile489, Ala491 and Arg510 in Bem1p are not conserved. Thus, these residues can be considered to be specificity determinants. In addition, the deletion of as few as five residues (335–339) from the C-terminus of p40^{phox} abolished the yeast two-hybrid interaction with p67^{phox} PB1 (Nakamura *et al.*, 1998). These residues are not included in the ubiquitin fold (237–329) (Figure 5; unpublished data), but are located on the C-terminal extension (330–339: DNYRVYNTMP) outside the

ubiquitin fold. Such an additional binding site can be a specificity determinant. On the other hand, some PB1 domains, such as aPKC PB1 and MEK5 PB1, can interact with several binding partners; therefore the selection of the binding partners and its biological significance are required to be determined for future studies.

Based on a homology search using the known PB1 family protein sequences as a template, nearly 200 PB1 family proteins including isoforms were found and are enrolled in the database SMART (Letunic *et al.*, 2002). Although most of their molecular functions are elusive, functional implications can be obtained by a search of the binding partners in the PB1 family. For example, mutational studies of a TFG PB1 portion of the TRK-T3 oncoprotein (a fused protein between an N-terminal portion of TFG and a tyrosine kinase portion of TrkA) were recently carried out (Roccato *et al.*, 2003); the NIH3T3 cells bearing a TRK-T3 mutant protein where a conserved lysine residue in the PB1 domain is substituted to an alanine residue (K14A) markedly reduced the relative transformation activity. This indicates that an unknown PB1 domain-containing binding partner of TFG PB1 has an important role in the transformation activity of the TRK-T3 oncoprotein.

Conclusions

This article describes the solution structure of Cdc24 PCCR (761–854). It is intriguing that Cdc24 PCCR (761–854) takes a ubiquitin fold despite the low sequence

Fig. 6. Classification of the PB1 family proteins based on the binding sites used for interaction with the cognate partners. The binding sites are mapped on (A) ribbon diagrams and (B) topology diagrams. The PB1 domain type I (Cdc24p) and type II (Bem1p) interact with each other using distinct binding sites on the ubiquitin fold, which are enclosed in magenta and green, respectively. The orientation of Cdc24 PB1 on the right of (A) and Bem1 PB1 on the left of (A) is the same as in Figure 4C.

homology and presents the PC motif portion on the ubiquitin scaffold. Mutational studies confirm that the PC motif portion of Cdc24 PCCR is a binding site for Bem1 PB1. Moreover, we could align the amino acid sequences of the PB1 domains and the PCCRs based on the 3D structures. Therefore, the PB1 domains and the PCCRs were grouped into a novel family, referred to as the PB1 family. The extended PB1 family proteins form specific dimers with each other using distinct binding sites on the ubiquitin fold. Because of its robustness, a ubiquitin

fold can be widely used as a ubiquitous scaffold to present the residues or motifs necessary for interaction with other molecules in various sites.

Materials and methods

Sample preparation

For the preparation of Cdc24 PCCR (761–854), the corresponding gene was amplified by PCR from yeast genomic DNA and then cloned into the *Bam*HI–*Eco*RI sites of a pGEX-6P vector (Amersham Pharmacia Biotech). Cdc24 PCCR (761–854) was expressed in *Escherichia coli* strain BL21 (DE3) cells with a glutathione *S*-transferase (GST) fused at the N-terminus. For an unlabeled protein, the transformed cells were grown in Luria–Bertani medium. We also prepared uniformly $^{13}\text{C}/^{15}\text{N}$ - and ^{15}N -labeled proteins by growing the cells in M9 medium containing $^{15}\text{NH}_4\text{Cl}$ (1 g/l; Shoko) and CELTONE-CN or CELTONE-N (1 g/l; Martek Biosciences) with $[\text{U-}^{13}\text{C}]\text{glucose}$ (2 g/l; Shoko) or unlabeled glucose. The GST fusion proteins were first purified by affinity chromatography using glutathione–Sepharose-4B resin (Amersham Pharmacia Biotech). After cleavage with PreScission protease (Amersham Pharmacia Biotech), the isolated proteins were purified by gel filtration on a Superdex 75 column (Amersham Pharmacia Biotech). Cdc24 PCCR (761–854) includes an extra Gly-Pro-Leu-Gly sequence at the N-terminus derived from the protease cleavage site. The identity and integrity of all proteins were confirmed by MALDI-TOF mass spectrometry and SDS-PAGE. The purified Cdc24 PCCR (761–854) proteins were concentrated to 1 mM for NMR studies in H_2O – D_2O (9:1) or D_2O buffers containing 50 mM potassium phosphate (pH 6.3), 150 mM NaCl and 1 mM NaN_3 .

NMR spectroscopy

NMR measurements of Cdc24 PCCR (761–854) were carried out at 25°C on Varian UNITYplus 600 or UNITYINOVA 500 spectrometers. The sequential assignments of ^1H , ^{13}C and ^{15}N of Cdc24 PCCR (761–854) were achieved mainly by through-bond heteronuclear correlations along the backbone and the side chains with the following ^1H , ^{13}C and ^{15}N triple resonance spectra: 3D $\text{HN}(\text{CO})\text{CA}$, 3D HNCA , 3D $\text{CBCA}(\text{CO})\text{NH}$, 3D HNCACB , 3D $\text{HBHA}(\text{CBCACO})\text{NH}$, 3D $\text{C}(\text{CO})\text{NH}$ (Cavanagh *et al.*, 1996) and 4D HCCH-TOCSY (Olejniczak *et al.*, 1992) recorded on the $^{13}\text{C}/^{15}\text{N}$ -labeled sample. Distance information between protons was obtained from 3D ^{15}N -edited and ^{13}C -edited NOESY (Cavanagh *et al.*, 1996) recorded with a mixing time of 75 ms on the $^{13}\text{C}/^{15}\text{N}$ -labeled sample. Slowly exchanging amide protons were identified from a series of 2D ^1H – ^{15}N HSQC spectra recorded after dissolving the lyophilized and protonated $^{13}\text{C}/^{15}\text{N}$ -labeled protein in D_2O .

Structure determination

The NMRPipe software suite (Delaglio *et al.*, 1995) was used to process the NMR data and to pick the resonance peaks. The Olivia (Objective, chemical shift library based and stereo view assignment) software developed in our laboratory (M.Yokochi and F.Inagaki, in preparation) was used to assign the resonance peaks semi-automatically. The NOE peak assignments and the structure calculations were carried out using the CNS program (Brünger *et al.*, 1998) with the ARIA module (Nilges *et al.*, 1997). The ϕ dihedral angle restraints of Cdc24 PCCR (761–854) were obtained from the chemical shift analysis using the program TALOS (Cornilescu *et al.*, 1999). The restraint range was extended to $\pm 25^\circ$ for the residues located in α -helices or $\pm 40^\circ$ for the residues located in β -strands. The distance constraints for the hydrogen bonds were introduced on the basis of the slowly exchanging amide protons and the characteristic interproton NOE patterns. The restraint range was 2.8–3.4 Å for N–O pairs and 1.8–2.4 Å for H–O pairs. The restraints are summarized in Table I for Cdc24 PCCR (761–854). After nine cycles of iterative NOE peak assignments and structure calculation, a total of 100 structures were obtained. The average coordinate was obtained by averaging the coordinates of the 20 lowest-energy structures. The coordinates of Cdc24 PCCR (761–854) have been deposited in the Protein Data Bank (accession code 1Q1O).

The figures for the superimposed 20 structures were generated using the program MOLMOL (Koradi *et al.*, 1996). The ribbon diagrams were generated using the program Molscript (Kraulis, 1991). The molecular surface representations were generated using the program GRASP (Nicholls *et al.*, 1991).

In vitro pull-down binding assay using purified proteins

The DNA fragments for Bem1 PB1 and Cdc24 PCCR were ligated to pMAL vectors (New England Biolabs). The DNA fragments encoding point mutants of Bem1 PB1 and Cdc24 PCCR were obtained by PCR-mediated site-directed mutagenesis. The PCR products were ligated to pGEX-6P vectors (Amersham Pharmacia Biotech). All the constructs were sequenced to confirm their identities. GST and maltose binding protein (MBP) fusion proteins were expressed in *E. coli* strain BL21 (DE3) cells and purified by glutathione–Sephadex-4B resin (Amersham Pharmacia Biotech) or amylose resin (New England Biolabs), respectively (Noda et al., 2001; Terasawa et al., 2001). Pull-down binding assays were performed as previously described (Noda et al., 2001; Terasawa et al., 2001). Briefly, a pair comprising a GST fusion protein and an MBP fusion protein was incubated in 500 µl of phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 8.1 mM Na₂HPO₄, 1.5 mM KH₂PO₄ pH 7.4) containing 0.5% Triton X-100, and precipitated by glutathione–Sephadex-4B resin. After washing six times with the same buffer containing 0.5% Triton X-100, bound proteins were eluted with 10 mM glutathione. The eluates were subjected to 10% SDS–PAGE and stained with Coomassie brilliant blue. To estimate the amounts of protein on the gel, densitometric analysis was carried out using the software NIH Image.

Acknowledgements

We are grateful to Dr Hiroaki Terasawa of the Graduate School of Pharmaceutical Sciences, University of Tokyo, Dr Yukiko Noda of the Medical Institute of Bioregulation, Kyushu University, and Mr Yoshinori Hirano, Dr Satoru Yuzawa and Dr Masataka Horiuchi of Hokkaido University Graduate School of Pharmaceutical Sciences for their technical advice and fruitful discussions. This work has been supported by CREST of Japan Science and Technology (JST) and by Grant-in-Aids for Scientific Research on Priority Areas and National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

References

Akimoto, K., Mizuno, K., Osada, S., Hirai, S., Tanuma, S., Suzuki, K. and Ohno, S. (1994) A new member of the third class in the protein kinase C family, PKC lambda, expressed dominantly in an undifferentiated mouse embryonal carcinoma cell line and also in many tissues and cells. *J. Biol. Chem.*, **269**, 12677–12683.

Brünger, A.T. et al. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

Butty, A.-C., Perrinjaquet, N., Petit, A., Jaquenoud, M., Segall, J.E., Hofmann, K., Zwahlen, C. and Peter, M. (2002) A positive feedback loop stabilizes the guanine–nucleotide exchange factor Cdc24 at sites of polarization. *EMBO J.*, **21**, 1565–1576.

Campbell, I.G. et al. (1994) A novel gene encoding a B-box protein within the BRCA1 region at 17q21.1. *Hum. Mol. Genet.*, **3**, 589–594.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G., III, and Skelton, N.J. (1996) *Protein NMR Spectroscopy*. Academic Press, San Diego, CA.

Chan, A.W., Hutchinson, E.G., Harris, D. and Thornton, J.M. (1993) Identification, classification and analysis of beta-bulges in proteins. *Protein Sci.*, **2**, 1574–1590.

Chang, E.C., Barr, M., Wang, Y., Jung, V., Xu, H.-P. and Wigler, M.H. (1994) Cooperative interaction of *S. pombe* proteins required for mating and morphogenesis. *Cell*, **79**, 131–141.

Chenevert, J., Corrado, K., Bender, A., Pringle, J. and Herskowitz, I. (1992) A yeast gene (*BEM1*) necessary for cell polarization whose product contains two SH3 domains. *Nature*, **356**, 77–79.

Chothia, C. and Lesk, A.M. (1986) The relation between the divergence of sequence and structure in proteins. *EMBO J.*, **5**, 823–826.

Cornilescu, G., Marquardt, J.L., Ottiger, M. and Bax, A. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.*, **120**, 6836–6837.

Cornilescu, G., Delaglio, F. and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX PIPES. *J. Biomol. NMR*, **6**, 277–293.

Devergne, O., Hummel, M., Koeppen, H., Le Beau, M.M., Nathanson, E.C., Kieff, E. and Birkenbach, M. (1996) A novel interleukin-12 p40-related

protein induced by latent Epstein–Barr virus infection in B lymphocytes. *J. Virol.*, **70**, 1143–1153.

English, J.M., Vanderbilt, C.A., Xu, S., Marcus, S. and Cobb, M.H. (1995) Isolation of MEK5 and differential expression of alternatively spliced forms. *J. Biol. Chem.*, **270**, 28897–28902.

Etienne-Manneville, S. and Hall, A. (2003) Cell polarity: Par6, aPKC and cytoskeletal crosstalk. *Curr. Opin. Cell Biol.*, **15**, 67–72.

Gao, L., Joberty, G. and Macara, I.G. (2002) Assembly of epithelial tight junctions is negatively regulated by Par6. *Curr. Biol.*, **12**, 221–225.

Greco, A., Mariani, C., Miranda, C., Lupas, A., Pagliardini, S., Pomati, M. and Pierotti, M.A. (1995) The DNA rearrangement that generates the *TRK-T3* oncogene involves a novel gene on chromosome 3 whose product has a potential coiled-coil domain. *Mol. Cell Biol.*, **15**, 6118–6127.

Hernández, L. et al. (1999) *TRK*-fused gene (TFG) is a new partner of *ALK* in anaplastic large cell lymphoma producing two structurally different *TFG-ALK* translocations. *Blood*, **94**, 3265–3268.

Hernández, L. et al. (2002) Diversity of genomic breakpoints in *TFG-ALK* translocations in anaplastic large cell lymphomas: identification of a new *TFG-ALK_{xl}* chimeric gene with transforming activity. *Am. J. Pathol.*, **160**, 1487–1494.

Hung, T.-J. and Kemphues, K.J. (1999) PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in *Caenorhabditis elegans* embryos. *Development*, **126**, 127–135.

Ito, T., Matsui, Y., Ago, T., Ota, K. and Sumimoto, H. (2001) Novel modular domain PB1 recognizes PC motif to mediate functional protein–protein interactions. *EMBO J.*, **20**, 3938–3946.

Joberty, G., Petersen, C., Gao, L. and Macara, I.G. (2000) The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42. *Nat. Cell Biol.*, **2**, 531–539.

Joung, I., Strominger, J.L. and Shin, J. (1996) Molecular cloning of a phosphotyrosine-independent ligand of the p56^{lck} SH2 domain. *Proc. Natl Acad. Sci. USA*, **93**, 5991–5995.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55.

Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. D*, **58**, 381–391.

Kuribayashi, F., Nunoi, H., Wakamatsu, K., Tsunawaki, S., Sato, K., Ito, T. and Sumimoto, H. (2002) The adaptor protein p40^{phox} as a positive regulator of the superoxide-producing phagocyte oxidase. *EMBO J.*, **21**, 6312–6320.

Lamark, T., Perander, M., Outzen, H., Kristiansen, K., Øvervatn, A., Michaelsen, E., Bjørkøy, G. and Johansen, T. (2003) Interaction codes within the family of mammalian PB1 domain-containing proteins. *J. Biol. Chem.*, **278**, in press.

Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.

Leto, T.L., Lomax, K.J., Volpp, B.D., Nunoi, H., Sechler, J.M., Nauseef, W.M., Clark, R.A., Gallin, J.I. and Malech, H.L. (1990) Cloning of a 67-kD neutrophil oxidase factor with similarity to a noncatalytic region of p60^{c-src}. *Science*, **248**, 727–730.

Letunic, I. et al. (2002) Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res.*, **30**, 242–244.

Marcus, S.L., Winrow, C.J., Capone, J.P. and Rachubinski, R.A. (1996) A p56^{lck} ligand serves as a coactivator of an orphan nuclear hormone receptor. *J. Biol. Chem.*, **271**, 27197–27200.

Mencinger, M., Panagopoulos, I., Andreasson, P., Lassen, C., Mitelman, F. and Aman, P. (1997) Characterization and chromosomal mapping of the human *TFG* gene involved in thyroid carcinoma. *Genomics*, **41**, 327–331.

Milner-White, E.J. (1987) Beta-bulges within loops as recurring features of protein structure. *Biochim. Biophys. Acta*, **911**, 261–265.

Miyamoto, S., Ohya, Y., Ohsumi, Y. and Anraku, Y. (1987) Nucleotide sequence of the CLS4 (CDC24) gene of *Saccharomyces cerevisiae*. *Gene*, **54**, 125–132.

Moscat, J. and Diaz-Meco, M.T. (2000) The atypical protein kinase Cs. Functional specificity mediated by specific protein adaptors. *EMBO Rep.*, **1**, 399–403.

Nakamura, R., Sumimoto, H., Mizuki, K., Hata, K., Ago, T., Kitajima, S., Takeshige, K., Sakaki, Y. and Ito, T. (1998) The PC motif: a novel and evolutionarily conserved sequence involved in interaction between p40^{phox} and p67^{phox}, SH3 domain-containing cytosolic factors of the phagocyte NADPH oxidase. *Eur. J. Biochem.*, **251**, 583–589.

- Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.
- Nilges,M., Macias,M.J., O'Donoghue,S.I. and Oschkinat,H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β -spectrin. *J. Mol. Biol.*, **269**, 408–422.
- Noda,Y., Takeya,R., Ohno,S., Naito,S., Ito,T. and Sumimoto,H. (2001) Human homologues of the *Caenorhabditis elegans* cell polarity protein PAR6 as an adaptor that links the small GTPases Rac and Cdc42 to atypical protein kinase C. *Genes Cells*, **6**, 107–119.
- Noda,Y., Kohjima,M., Izaki,T., Ota,K., Yoshinaga,S., Inagaki,F., Ito,T. and Sumimoto,H. (2003) Molecular recognition in dimerization between PB1 domains. *J. Biol. Chem.*, **278**, in press.
- Ohno,S. (2001) Intercellular junctions and cellular polarity: the PAR-aPKC complex, a conserved core cassette playing fundamental roles in cell polarity. *Curr. Opin. Cell Biol.*, **13**, 641–648.
- Olejniczak,E.T., Xu,R.X. and Fesik,S.W. (1992) A 4D HCCH-TOCSY experiment for assigning the side chain ^1H and ^{13}C resonances of proteins. *J. Biomol. NMR*, **2**, 655–659.
- Ono,Y., Fujii,T., Ogita,K., Kikkawa,U., Igarashi,K. and Nishizuka,Y. (1989) Protein kinase C zeta subspecies from rat brain: its structure, expression and properties. *Proc. Natl Acad. Sci. USA*, **86**, 3099–3103.
- Ponting,C.P. (1996) Novel domains in NADPH oxidase subunits, sorting nexins and PtdIns 3-kinases: binding partners of SH3 domains? *Protein Sci.*, **5**, 2353–2357.
- Ponting,C.P., Ito,T., Moscat,J., Diaz-Meco,M.T., Inagaki,F. and Sumimoto,H. (2002) OPR, PC and AID: all in the PB1 family. *Trends Biochem. Sci.*, **27**, 10.
- Puls,A., Schmidt,S., Grawe,F. and Stabel,S. (1997) Interaction of protein kinase C ζ with ZIP, a novel protein kinase C-binding protein. *Proc. Natl Acad. Sci. USA*, **94**, 6191–6196.
- Roccatto,E., Pagliardini,S., Cleris,L., Canevari,S., Formelli,F., Pierotti,M.A. and Greco,A. (2003) Role of TFG sequences outside the coiled-coil domain in TRK-T3 oncogenic activation. *Oncogene*, **22**, 807–818.
- Selbie,L.A., Schmitz-Peiffer,C., Sheng,Y. and Biden,T.J. (1993) Molecular cloning and characterization of PKC iota, an atypical isoform of protein kinase C derived from insulin-secreting cells. *J. Biol. Chem.*, **268**, 24296–24302.
- Sibanda,B.L. and Thornton,J.M. (1985) β -hairpin families in globular proteins. *Nature*, **316**, 170–174.
- Sibanda,B.L., Blundell,T.L. and Thornton,J.M. (1989) Conformation of β -hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J. Mol. Biol.*, **206**, 759–777.
- Terasawa,H., Noda,Y., Ito,T., Hatanaka,H., Ichikawa,S., Ogura,K., Sumimoto,H. and Inagaki,F. (2001) Structure and ligand recognition of the PB1 domain: a novel protein module binding to the PC motif. *EMBO J.*, **20**, 3947–3956.
- Watts,J.L., Etemad-Moghadam,B., Guo,S., Boyd,L., Draper,B.W., Mello,C.C., Priess,J.R. and Kemphues,K.J. (1996) *par-6*, a gene involved in the establishment of asymmetry in early *C.elegans* embryos, mediates the asymmetric localization of PAR-3. *Development*, **122**, 3133–3140.
- Whitehouse,C., Chambers,J., Howe,K., Cobourne,M., Sharpe,P. and Solomon,E. (2002) NBR1 interacts with fasciculation and elongation protein zeta-1 (FEZ1) and calcium and integrin binding protein (CIB) and shows developmentally restricted expression in the neural tube. *Eur. J. Biochem.*, **269**, 538–545.
- Wientjes,F.B., Hsuan,J.J., Totty,N.F. and Segal,A.W. (1993) p40^{phox}, a third cytosolic component of the activation complex of the NADPH oxidase to contain *src* homology 3 domains. *Biochem. J.*, **296**, 557–561.
- Zhou,G., Bao,Z.Q. and Dixon,J.E. (1995) Components of a new human protein kinase signal transduction pathway. *J. Biol. Chem.*, **270**, 12665–12669.

Received June 4, 2003; revised July 29, 2003;
accepted July 30, 2003