Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The crystal structure of PEBP-2, a homologue of the PEBP/RKIP family

P. C. Simister, M. J. Banfield and R. L. Brady*

Department of Biochemistry, University of Bristol, Bristol BS8 1TD, England

Correspondence e-mail: l.brady@bristol.ac.uk

Received 22 December 2001
Accepted 21 March 2002

PDB Reference: mPEBP2, 1kn3, r1kn3.

Proteins from the PEBP (phosphatidylethanolamine-binding protein) family have been identified in a wide variety of species and are thought to regulate a range of intracellular signalling cascades. The rat homologue (known as RKIP; Raf-1 kinase inhibitor protein) has been shown to negatively regulate the MAP kinase pathway through formation of inhibitory complexes with Raf-1 and MEK. The crystal structure of a new, murine member of the PEBP family, termed mPEBP-2, has been determined. On the basis of amino-acid homology, mPEBP-2 belongs to a distinct subset of the mammalian PEBP proteins. Nonetheless, mPEBP-2 is seen to be very similar in structure to other PEBP proteins from human, bovine and plant sources. Regions of distinctive sequence associated with the PEBP-2 subset are discussed with reference to this structure.

1. Introduction

The phosphatidylethanolamine-binding protein (PEBP) family is a highly conserved group of proteins that has homologues in a wide variety of organisms. Previously, a number of functions have been suggested for the mammalian proteins that include lipid binding (Schoentgen & Jolles, 1995), inhibition of serine proteases (Hengst *et al.*, 2001) and that the protein is a precursor for a bioactive peptide (HCNP) important in development of the hippocampus (Tohdoh *et al.*, 1995). The plant PEBP homologues are involved in the control of a morphogenic switch between shoot growth and flower structures (infloresecence; Bradley *et al.*, 1996). Recently, it was demonstrated that rat PEBP was able to specifically interfere with MEK phosphorylation and activation by Raf-1 kinase (Yeung *et al.*, 1999), a crucial step in intracellular signalling pathways controlling mitogenesis and cell differentiation (in this study, the protein was termed RKIP; Raf-1 kinase inhibitor protein). These experiments suggested that inhibition results from a conformational change in Raf-1 associated with binding of PEBP/RKIP or through direct steric hindrance of the Raf-1–MEK interaction (Yeung *et al.*, 2000). RKIP also appears to specifically bind to MEK (Yeung *et al.*, 2000). As the plant homologues also regulate cell development, it has been suggested that they also act *via* interactions with cellular kinases. A serine-threonine kinase has been identified as a molecular ligand for the tomato PEBP homologue (E. Lifschitz & L. Pneuli, personal communication).

As the number of genomes being sequenced increases and further studies on this protein family are undertaken, it has become apparent that many organisms contain several forms of PEBP. For instance, protein sequence databases for the plants *Arabidopsis thaliana* and *Oryza sativa* (rice) each show six sequences belonging to the PEBP family. This trend also extends into higher organisms, where the fruit fly (*Drosophila melogaster*) genome is now know to contain at least five PEBP paralogues. Recently, new forms of mammalian PEBPs have also been isolated from cDNA libraries from both mice and rats during the process of screening for novel cDNAs involved in spermiogenesis (O'Bryan, 2002). These proteins, which share a sequence identity of 91%, have been named rat and mouse PEBP-2 (rPEBP-2 and mPEBP-2, respectively). These proteins form a distinct sequence subset of the mammalian PEBP family, with mPEBP-2, for example, sharing 84% sequence identity with human PEBP-1 and 79% identity with the first murine PEBP identified (mPEBP-3). No PEBP-2-like homologue has yet been identified in humans.

Forms of mPEBP and hPEBP mRNA are expressed in many adult tissues (Seddiqi *et al.*, 1994). In contrast to this, mPEBP-2 RNA is exclusively expressed in testis (O'Bryan, 2002), supporting the notion that mPEBP-2 plays a specific role in spermiogenesis. Recently, evidence has emerged that activation of the MAP-kinase pathway occurs during spermiogenesis and post-testis sperm maturation (*e.g.* Berruti, 2000). mPEBP-2 has been shown to bind both B-Raf and MEK1 and displays an

overlapping expression pattern with components of the MAP kinase signalling pathway during sperm maturation (O'Bryan, 2002). It is therefore possible that mPEBP-2 may exert control over sperm development by specifically regulating the MAP-kinase pathway. mPEBP-3 mRNA is also expressed in the testis, suggesting there may be a role for both forms of the PEBP family, possibly controlling MAP-kinase signalling in a synergistic fashion.

To further probe the role of these different forms of PEBP from a single organism, in this study we have analysed the known sequences for mammalian PEBPs. We have also expressed, crystallized and determined the crystal structure of mPEBP-2 and compare its structure with those previously determined for the PEBP family.

2. Materials and methods

2.1. Protein production

A plasmid containing the DNA encoding mPEBP-2 was provided by Dr Moira O'Bryan (Monash Institute of Reproduction and Development, Australia). Although this clone lacks the four N-terminal residues of the protein, these amino acids are consistently disordered in all crystal structures that have been determined for the PEBP family and the absence of these residues is not expected to affect the structure of the protein. The mPEBP-2 DNA was cloned into the pET28a vector (Novagen) for high-level expression in *Escherichia coli*, using standard techniques.

The protein was expressed and purified as described for hPEBP-1 (Banfield *et al.*, 1998), with the final protein solution comprising 7.5 mg ml^{-1} mPEBP-2 in 20 m*M* bis–Tris, 100 m*M* NaCl pH 6.

2.2. Crystallization and data collection

Diffraction-quality crystals of mPEBP-2 were obtained from a 1:1 mixture of protein solution with 30% PEG 4000, 200 m*M* sodium acetate trihydrate buffered with 100 m*M* Tris–HCl pH 8.5. Prior to data collection, crystals were cryopreserved by soaking in a solution comprising the precipitant as above supplemented with 5% glycerol. Crystals were then frozen at 100 K in a liquid-nitrogen cold stream. X-ray data were collected on station PX14.1 at the Daresbury SRS to 1.8 Å resolution. The crystals grew as very thin plates, resulting in significant differences in overall diffraction intensity and changes in mosaicity at different crystal orientations. To account for this, the data set was divided into batches of 20 frames for processing. The data were processed and scaled with the *HKL* suite (Otwinowski & Minor, 1996) and a summary of both data-collection parameters and crystal properties is shown in Table 1.

2.3. Structure solution and refinement

The structure of mPEBP-2 was solved by molecular replacement using *AMoRe* (Navaza, 1994) as implemented in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). A monomer of hPEBP-1 (Banfield *et al.*, 1998; PDB code 1bd9) was used as a search model. A single solution was returned from the rotation and translation functions, as expected from the volume of the unit cell. The structure of mPEBP-2 was refined with *REFMAC*5 (Murshudov *et al.*, 1997) and all model building was performed with *O* (Jones *et al.*, 1991). Repeated cycles of refinement/rebuilding lowered $R_{\rm cryst}$ and $R_{\rm free}$ (Brünger, 1992) to final values of 21.0 and 26.7%, with root-mean-square deviations (r.m.s.d.s) of 0.012 Å for bond lengths and 1.49° for bond

Table 1
Data-collection and refinement parameters for mPEBP2.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters (Å)	a = 40.82, b = 51.40, c = 79.77
Space group	$P2_12_12_1$
Wavelength (Å)	1.488
Resolution range (Å)	30–1.8
No. of unique reflections	15638
Redundancy	4.7 (3.3)
Completeness (%)	96.4 (92.4)
$R_{\rm merge}$ (%)	8.4 (11.6)
$I/\sigma(I)$	12.9 (8.5)
Refinement	
Resolution range (Å)	20–1.8 (1.88–1.8)
$R_{\rm cryst}$ (%)	21.0 (22.0)
$R_{\rm free}$ (%)	26.7 (24.0)
R.m.s.d., bond lengths (Å)	0.012
R.m.s.d., bond angles (°)	1.49
No. of non-H protein atoms	11744
No. of water molecules	124
Average B value (Å^2)	
Main chain	11.4
Side chain	12.0
Solvent molecules	29.2

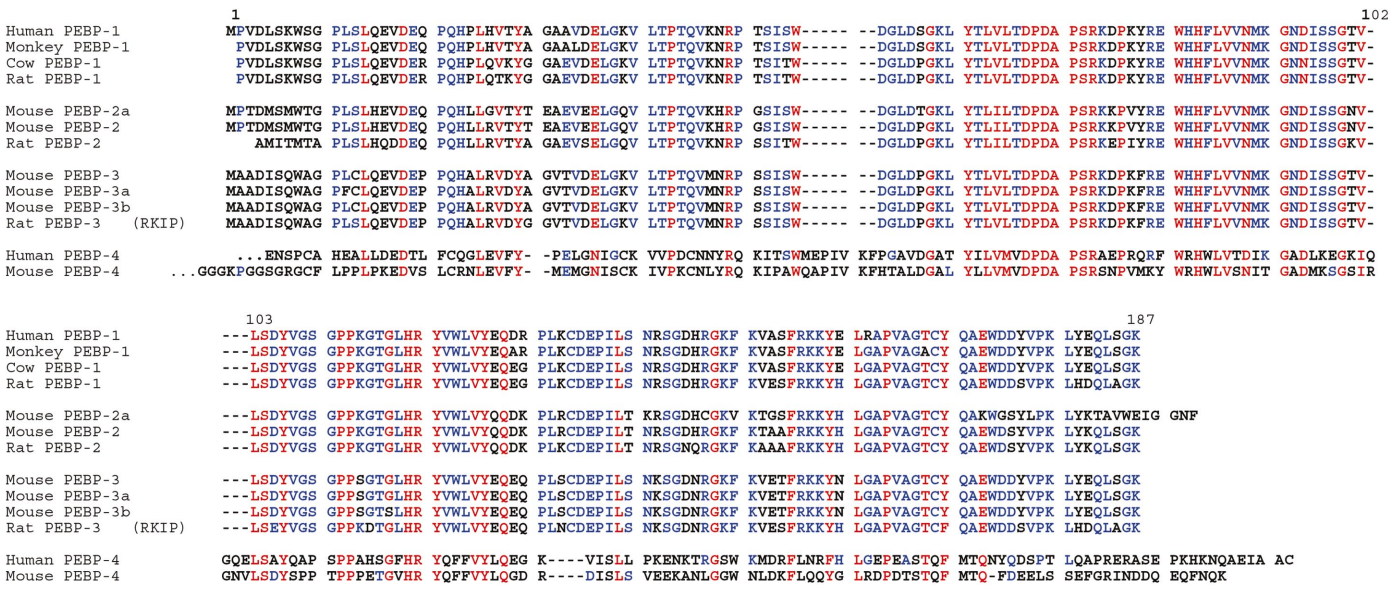

Figure 1
Sequence alignment of all available mammalian PEBP proteins, prepared with *MULTALIN* (Corpet, 1988). The proteins are divided into four families according to sequence homology. The first 33 and 40 N-terminal residues for hPEBP-4 and mPEBP-4, respectively, are not shown in the figure.

angles. The Ramachandran plot (Ramachandran & Sasisekharan, 1968) for the final model shows 100% of residues lie in 'core' regions as described in Kleywegt & Jones (1996).

3. Results and discussion

3.1. Sequence analysis

At present, there are 13 identified mammalian PEBP sequences. An alignment of these protein sequences is shown in Fig. 1 and suggests that these proteins can be grouped into four subfamilies (here termed PEBP-1, PEBP-2, PEBP-3 and PEBP-4). Three of these subfamilies have previously been noted (O'Bryan, 2002). The majority of differences in protein sequence between the subfamilies are observed in the N-terminal 40 residues, with most variation displayed within the first ten amino acids. Members of the first three subfamilies share similar overall features, each being approximately 190 residues in length, with no insertions or deletions. The main features that distinguish the fourth family (excluding the extension at the N-terminus) are the presence of two insertions and one deletion in the protein sequence. These two insertions (between residues 55 and 56, and 102 and 103 in hPEBP-1 numbering) and single deletion (comprising residues 131–134 in hPEBP-1) are all located in loop regions as identified from PEBP crystal structures. These insertions and deletions are not expected to affect the overall fold of the protein.

3.2. Overall structure

The final model of mPEBP-2 spans residues Ser6–Ser185 of the native protein; 124 solvent sites have been modelled as water molecules. The overall fold of mPEBP-2 is very similar to previously determined mammalian PEBP structures (Banfield *et al.*, 1998; Serre *et al.*, 1998), comprising a central β-sheet flanked by a second, smaller, β-sheet on one side and an α-helix on the other. The protein also shares the overall PEBP fold with plant (Banfield & Brady, 2000) and bacterial forms (Serre *et al.*, 2001). mPEBP-2 overlays on hPEBP-1 with an r.m.s.d. of 0.27 Å (178 equivalent C^{α} atoms), reflecting the extensive level of fold conservation evident throughout the structure (see Fig. 2*a*). Small deviations are apparent at the N-terminus and in some loop regions.

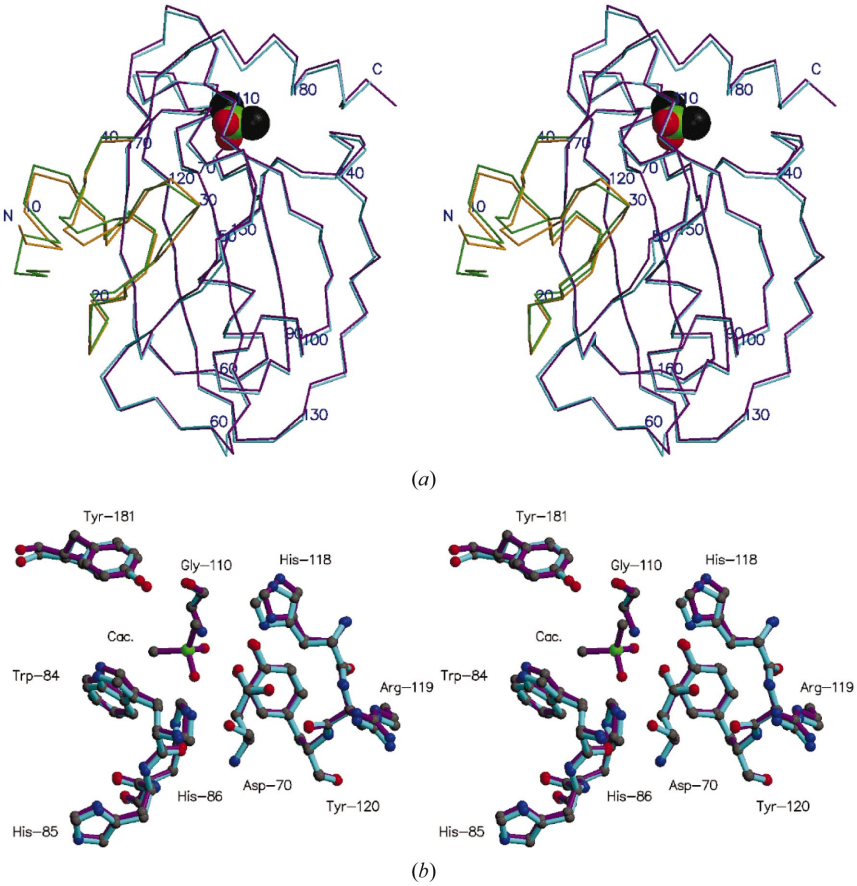

Figure 2
Stereoviews showing overlays of the mPEBP-2 and hPEBP-1 structures. (*a*) C^{α} trace showing an overlay of the mPEBP-2 and hPEBP-1 structures. The first 41 residues are coloured orange (mPEBP-2) and green (hPEBP-1), highlighting the surface region where the greatest sequence differences in the mammalian PEBP family are clustered. The rest of the C^{α} trace is coloured cyan (mPEBP-2) and purple (hPEBP-1). The ligand-binding site is identified by the bound cacodylate ion shown in CPK (from the hPEBP-1 structure). (*b*) Amino acids in the ligand-binding site [mPEBP-2 bonds in cyan, hPEBP-1 bonds in purple; bound cacodylate (from the hPEBP-1 structure) is also shown].

3.3. Functional sites on mPEBP-2

From sequence analysis and existing PEBP structures, a number of regions have been identified that are thought to be functionally important. Many of these are close to the identified ligand-binding site or are believed to be essential for maintaining the architecture of this region. These include a DPD*x*P*x*$_n$H motif (residues 69–86, where *n* is 11 in all mammalian proteins and the second proline adopts a *cis*-peptide conformation), a G*x*HR motif (residues 116–119) and a non-prolyl *cis*-peptide bond conformation adopted by residue 83. Unusually, neither of the *cis*-peptide conformations is observed in the bacterial PEBP structures. The crystal structure of mPEBP-2 supports the expectation that the anionic ligand-binding site is central to the function of PEBP proteins. The DPD*x*P*x*$_n$H and G*x*HR motifs are conserved in mPEBP-2 at both the sequence and structural levels, as are both of the *cis*-peptide conformations. The ligand-binding site of mPEBP-2 can be overlaid with a very high degree of conservation on any of the mammalian or plant structures determined to date (Fig. 2*b*). In the mPEBP-2 structure, this site is occupied by a network of water molecules, as was observed in the structure of the plant PEBP homologue (Banfield & Brady, 2000). The observed electron density is not consistent with any other type of ligand. The bound solvent at this site suggests it may be feasible to introduce other ligands into the crystal lattice in future studies.

Another region of interest is the N-terminus, part of which appears to be cleaved from mammalian forms of these protein to release the bioactive peptide HCNP (hippocampal neurostimulatory peptide). Relative to its PEBP-3 subfamily homologue, mPEBP-2 has four substitutions in the first ten amino-terminal residues. Sequence variation between subgroups of PEBP proteins is most evident within the first N-terminal 40 residues and much of this sequence maps to a single region of the protein surface. This region is some distance

from the ligand-binding site (Fig. 2*a*). It is not clear whether this variation represents an adaptation to confer specificity between the various PEBPs and different interacting proteins and/or regulate affinity or if variation in this region is permitted as it is remote from the functional protein–protein interface. The picture of regulation by PEBP subgroups is complex. For instance, in *Arabidopsis* two paralogues have been shown to have the same function but differ in their tissue distribution (TFL1 and ATC; Mimida *et al.*, 2001), implying that a conserved functional surface might be expected across PEBP subgroups from a single species. However, also in *Arabidopsis*, another paralogue (FT) acts as an antagonist to TFL1 (Kobayashi *et al.*, 1999). Whether this implies interaction with the same target protein but promoting an opposite effect (potentially involving a conserved functional surface) or interaction with a different protein (suggesting a different functional surface) is yet to be determined. Further studies of these proteins, in particular of their complexes with interacting proteins, are essential in order to elucidate the significance of the variations observed between PEBP subtypes.

The authors thank Dr Moira O'Bryan and coworkers (Monash Institute of Reproduction and Development, Australia) for their gift of the DNA encoding mPEBP-2 and for discussions. We also thank the Human Frontiers in Science Program for funding PEBP studies, and the staff of the CCLRC Daresbury SRS Laboratory for their help with data collection.

References

Banfield, M. J., Barker, J. J., Perry, A. C. F. & Brady, R. L. (1998). *Structure*, **6**, 1245–54.
Banfield, M. J. & Brady, R. L. (2000). *J. Mol. Biol.* **297**, 1159–1170
Berruti, G. (2000). *Exp. Cell Res.* **257**, 172–179.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Corpet, F. (1988). *Nucleic Acids Res.* **16**, 10881–10890.
Bradley, D., Carpenter, R., Copsey, L., Vincent, C., Rothstein, S. & Coen, E. (1996). *Nature (London)*, **379**, 791–797.
Hengst, U., Albrecht, H., Hess, D. & Monard, D. (2001). *J. Biol. Chem.* **276**, 535–540.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.
Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M. & Araki, T. (1999). *Science*, **286**, 1960–1962.
Mimida, N., Goto, K., Kabayashi, Y., Araki, T., Ahn, J. H., Weigel, D., Murata, M., Motoyoshi, F. & Sakamoto, W. (2001). *Genes Cells*, **6**, 327–336.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
O'Bryan, M. (2002). Submitted.
Ramachandran & Sasisekharan (1968). *Adv. Protein Chem.* **23**, 283–437.
Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.
Schoentgen, F. & Jolles, P. (1995). *FEBS Lett.* **369**, 22–26.
Seddiqi, N., Bollengier, F., Alliel, P. M., Perin, J. P., Bonnet, F., Bucquoy, S., Jolles, P. & Schoentgen, F. (1994). *J. Mol. Evol.* **39**, 655–660.
Serre, L., Pereira de Jesus, K., Zelwer, C., Bureaud, N., Schoentgen, F. & Bénédetti, H. (2001). *J. Mol. Biol.* **310**, 617–634.
Serre, L., Vallee, B., Bureaud, N., Schoentgen, F. & Zelwer, C. (1998). *Structure*, **6**, 1255–1265.
Tohdoh, N., Tojo, S., Agui, H. & Ojika, K. (1995). *Mol. Brain Res.* **30**, 381–384.
Yeung, K., Janosch, P., McFerran, B., Rose, D. W., Mischak, H., Sedivy, J. M. & Kolch, W. (2000). *Mol. Cell. Biol.* **20**, 3079–3085.
Yeung, K., Seitz, T., Li, S. F., Janosch, P., McFerran, B., Kaiser, C., Fee, F., Katsanakis, K. D., Rose, D. W., Mischak, H., Sedivy, J. M. & Kolch, W. (1999). *Nature (London)*, **401**, 173–177.