Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, characterization and crystallization in two crystal forms of bovine cyclophilin 40

Jacqueline Dornan,[a] Paul Taylor,[a] Amerigo Carrello,[b,c] Rodney F. Minchin,[c] Thomas Ratajczak[b,c] and Malcolm D. Walkinshaw[a]*

[a]Structural Biochemistry Group, Institute of Cell and Molecular Biology, The University of Edinburgh, Michael Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, Scotland, [b]Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Hospital Avenue, Nedlands WA 6009, Australia, and [c]Department of Pharmacology, University of Western Australia, The Queen Elizabeth II Medical Centre, Nedlands WA 6009, Australia

Correspondence e-mail: m.walkinshaw@ed.ac.uk

Received 7 December 1998
Accepted 22 February 1999

The purification and crystallization of two different crystal forms of the two-domain protein bovine cyclophilin 40 is reported. Tetragonal crystals grown in methyl pentanediol belong to space group $P4_222$ with unit-cell parameters a = 94.5, c = 118.3 Å. Long thin needles grown from PEG belong to space group $C2$ with unit-cell parameters a = 125.71, b = 47.3, c = 74.6 Å, β = 93.90°. The N-terminal 170 amino acids have significant homology with the well characterized human cyclophilin A. The C-terminal domain is largely made up of three copies of the tetratricopeptide repeat motif thought to be involved in mediating protein–protein interactions. Cyclophilins are frequently found as domains in larger multidomain proteins. To date, only X-ray structures of single-domain cyclophilins have been reported, and this work provides the first example of the purification and crystallization of a larger protein containing a cyclophilin domain.

1. Introduction

Cyclophilins are phylogenetically highly conserved proteins which act as peptidyl–prolyl *cis–trans* isomerases (PPIases) and are involved in protein folding/refolding as well as protein–protein interactions (Galat & Metcalfe, 1995). The first member of this expanding family to be discovered was the 18 kDa human cyclophilin A (hCyP18) which was identified by its binding to the fungal peptide immunosuppressant cyclosporin A (CsA) (Handschumacher *et al.*, 1984). The X-ray structures of cyclophilin A and of its complex with CsA have been determined (Kallen *et al.*, 1991; Pflugl *et al.*, 1993) and show the eight-stranded barrel structure conserved in all cyclophilin domains.

Larger multidomain proteins such as CyP150 from natural killer cells (Anderson *et al.*, 1993) and RanBP2 (Yokoyama *et al.*, 1995) also contain single cyclophilin domains. The two-domain cyclophilin 40 (CyP40) protein family has been found in mammals (Kieffer *et al.*, 1992 1993; Ratajczak *et al.*, 1993) and yeast (Weisman *et al.*, 1996; Duina *et al.*, 1996), and was originally isolated and characterized in association with the unactivated non-DNA binding form of the oestrogen receptor (Ratajczak *et al.*, 1993). CyP40 is also present as one of several components of mature unactivated complexes of the glucocorticoid and progesterone receptors (reviewed in Pratt & Toft, 1997). The major chaperones, hsp90 and hsp70, are included as non-hormone-binding components within these steroid-receptor complexes, together with p23 and the immunophilins FKBP51 and FKBP52 (Pratt & Toft, 1997), which are cellular targets of the immunosuppressant FK506 (Galat & Metcalfe, 1995). CyP40, FKBP51 and FKBP52 belong to a distinct class of proteins which display an N-terminal immunophilin-like domain with overlapping regions for immunosuppressant drug interaction and isomerase activity, together with a conserved C-terminal tetratricopeptide repeat (TPR) domain proposed to mediate protein–protein interaction (Ratajczak *et al.*, 1993; Nair *et al.*, 1997). Through this conserved C-terminal region, these immunophilins target a common interaction site in hsp90, resulting in the formation of separate immunophilin–hsp90–steroid-receptor complexes (Nair *et al.*, 1997; Radanyi *et al.*, 1994; Hoffmann & Handschumacher, 1995; Ratajczak & Carrello, 1996; Young *et al.*, 1998; Carrello *et al.*, 1999). CyP40 and its partner immunophilins appear to have a coordinate role with hsp90 in the functional control of receptor activity (Pratt & Toft, 1997).

In vitro folding assays have shown that Cyp40 and FKBP52 can function as molecular chaperones, similar to hsp90 and hsp70, by holding substrate proteins in a partially folded conformation (Freeman *et al.*, 1996; Bose *et al.*, 1996). The expression of these immunophilins is ubiquitous (Ratajczak *et al.*, 1993; Kieffer *et al.*, 1993; Nair *et al.*, 1997) and there is evidence for their association with hsp90 independently of steroid receptors (Pratt & Toft, 1997). Cpr6 and Cpr7 are the only *Saccharomyces cerevisiae* homologues of Cyp40, and both associate with hsp90 (Duina *et al.*, 1996). Mutations which decrease the level of hsp90 expression, coupled with a deletion of Cpr7, cause a

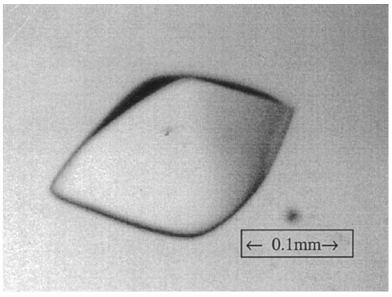
© 1999 International Union of Crystallography
Printed in Denmark – all rights reserved

significant reduction in the rate of cell growth (Duina *et al.*, 1996). The involvement of wis2, the corresponding CyP40 homolog in *Schizosaccharomyces pombe*, in regulating the G2-mitosis transition (Weisman *et al.*, 1996) is consistent with this observation. Hsp90 in association with Cpr7 has also been shown to participate in the negative regulation of the heat-shock response in *S. cerevisiae* (Duina *et al.*, 1998). Taken together, these results suggest that Cyp40 plays a major role in hsp90-mediated signal-transduction pathways.

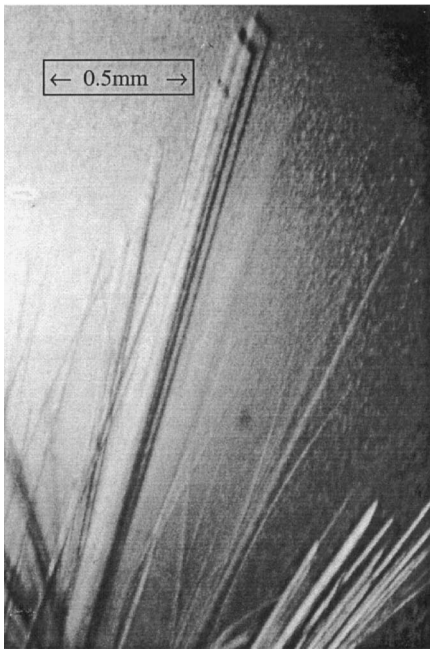
2. Materials and methods

2.1. Expression, purification and characterization

Expression plasmids for untagged wild-type bovine CyP40 (bCyP40) and the cyclophilin fused to glutathione S-transferase (GST) have been described (Ratajczak & Carrello, 1996; Carrello *et al.*, 1999).



(*a*)



(*b*)

Figure 1
Pictures of the monoclinic and tetragonal crystal forms of Cyp40.

GST-fused recombinant cyclophilin 40 (GST-bCyP40 WT; Ratajczak & Carrello, 1996) was expressed following a 4 h induction with 0.4 m*M* IPTG (ICN/FLOW). Cell pellets were frozen at 253 K, thawed slowly on ice and solubilized in lysis buffer [20 m*M* Tris buffer pH 7.8 containing 150 m*M* NaCl, 2 m*M* EDTA, 5 m*M* DTT, 1%(*v*/*v*) Triton X-100 and 5 m*M* benzamidine]. Lysozyme (ICN/FLOW) was added to 0.1%(*w*/*v*) and the cell suspension incubated on ice for 1 h. An equal volume of distilled water at 277 K was added and the cell suspension was mixed and centrifuged at 48 400*g* for 60 min. The resulting supernatant was frozen at 253 K, thawed on ice and filtered (0.45 µm) prior to purification. The protein extract was purified to near homogeneity using four purification steps. (i) Glutathione Sepharose (Pharmacia) affinity chromatography: protein was applied in buffer *A* (50 m*M* Tris–HCl pH 7.5 buffer) and GST-tagged protein was eluted with buffer *A* containing reduced 25 m*M* glutathione (Sigma). (ii) Anion-exchange chromatography: protein from the glutathione Sepharose chromatography was applied to Resource-Q resin (Pharmacia) in buffer *A*. Elution with a linear gradient from buffer *A* to buffer *A* plus 1 *M* NaCl gave peak fractions which were pooled, concentrated, quantitated for protein concentration by the Bradford method and dialysed overnight against buffer *A*. (iii) Thrombin cleavage: GST fusion protein was cleaved with purified thrombin at a ratio of 1:100 for 90 min at room temperature. (iv) Cleaved protein was subjected to a second round of chromatography on Resource-Q resin as already described, followed by a final affinity chromatography step on glutathione Sepharose.

Induction and harvesting of the non-fusion version of the protein was essentially as described above for the GST-bCyp40 WT protein. Purification to near homogeneity was achieved in two ion-exchange steps. The cell extract was applied to Q Sepharose (Amersham Pharmacia Biotech) in buffer *A*, the column was washed to baseline and then developed with a gradient of buffer *A* plus 1 *M* NaCl. Protein was eluted with 150 m*M* NaCl and fractions were analysed on 12% SDS–PAGE gels followed by staining with Coomassie Blue. Fractions containing bCyP40 were pooled and dialysed overnight at 277 K against 20 m*M* Tris pH 8.0 buffer (buffer *B*). The protein sample was applied to Mono-Q resin (PAB) in buffer *B*, and retained protein was eluted with a gradient of buffer *B* plus 0.5 *M* NaCl. Positive fractions were pooled and concentrated. Dynamic light-scattering analysis of the purified protein (DynaPro-801 with Microsampler attatchment, Protein Solutions, Inc.) showed it to be essentially monodisperse and stable, with an estimated molecular weight of 43 kDa.

The protein was further characterized by electrospray ionization mass spectrometry on a Micromass Platform II spectrometer. The sample was introduced into the instrument stream *via* an infusion pump and the *M*/*Z* spectrum was deconvoluted using the *MaxEnt* algorithm.

2.2. Crystallization

The hanging-drop vapour-diffusion method was used to grow crystals in 24-well Linbro plates. In all trials, crystals could only be obtained at 277 K.

Tetragonal crystals were grown by equilibrating a 4 µl drop of protein solution against 1 ml of well solution. The well solution consisted of 470–490 µl MPD, 100 µl 1 *M* imidazole pH 6.5, 200–360 µl 50%(*v*/*v*) glycerol solution and 50–230 µl water. The drop consisted of 2 µl protein solution and 2 µl well solution. The protein solution consisted of protein at between 10 and 80 mg ml^{-1} in 20 m*M* Tris buffer pH 8 containing 100 m*M* NaCl.

The morphology of the crystals in most drops was poorly defined and teardrop-shaped crystals grew with dimensions of ~0.05 mm. Addition of 1% glycerol and 5–10 m*M* DTT were found to improve crystal size and morphology (Fig. 1*a*), with crystals growing to 0.2 × 0.15 × 0.15 mm. The unit-cell dimensions are a = 94.5, c = 118.3 Å. A complete data set to a resolution of 2.6 Å has been collected on an Enraf–Nonius rotating-anode X-ray generator with a MAR Research 300 image plate.

Monoclinic needles were obtained using 12%(*v*/*v*) PEG as precipitant. The 1 ml well solution was composed of 240 µl of a 50% solution of PEG 2000, 100 µl 1 *M* cacodylate, 50 µl 1 *M* HCl, 200 µl 50% glycerol solution and 410 µl water. The 4 µl hanging drop consisted of 2 µl well solution plus 2 µl protein at 10–80 mg ml^{-1} in 20 m*M* Tris pH 8 buffer containing 100 m*M* NaCl. Very long thin plate-like needles grew within hours with typical dimensions 2 × 0.01 × 0.005 mm. These crystals are monoclinic with space group *C*2. Unit-cell parameters are a = 125.71, b = 47.3, c = 74.6 Å, β = 93.90°. These thin needle crystals typically diffract to a resolution of 3.6 Å on a rotating-anode X-ray source, and a complete data set has been collected.

```
            1                                                          50
CYPH_HUMAN  .......... ...VNPTVFF DIAVDGEPLG RVSFELFADK VPKTAENFRA
CYP3_CAEEL  .......... ..MSRSKVFF DITIGGKASG RIVMELYDDV VPKTAGNFRA
CYP4_HUMAN  MSHPSPQAKP SNPSNPRVFF DVDIGGERVG RIVLELFADI VPKTAENFRA
CYP4_BOVIN  MSHPSPQAKP SNPSNPRVFF DVDIGGERVG RIVLELFADI VPKTAENFRA
 Consensus  .......... .......vff d....g...g r...el..d. vpkta.nfra

            51                                                         100
CYPH_HUMAN  LSTGEKGFG. .......YKG SCFHRIIPGF MCQGGDFTRH NGTGGKSIYG
CYP3_CAEEL  LCTGENGIGK S.GKPLHFKG SKFHRIIPNF MIQGGDFTRG NGTGGESIYG
CYP4_HUMAN  LCTGEKGIGH TTGKPLHFKG CPFHRIIKKF MIQGGDFSNQ NGTGGESIYG
CYP4_BOVIN  LCTGEKGIGP TTGKPLHFKG CPFHRIIKKF MIQGGDFSNQ NGTGGESIYG
 Consensus  l.tge.g.g. ........kg ..fhrii..f m.qggdf... ngtgg.siyg

            101                                                        150
CYPH_HUMAN  EKFEDENFIL KHTGPGILSM ANAGPNTNGS QFFICTAKTE WLDGKHVVFG
CYP3_CAEEL  EKFPDENFKE KHTGPGVLSM ANAGPNTNGS QFFLCTVKTE WLDGKHVVFG
CYP4_HUMAN  EKFEDENFHY KHDREGLLSM ANAGRNTNGS QFFITTVPTP HLDGKHVVFG
CYP4_BOVIN  EKFEDENFHY KHDKEGLLSM ANAGSNTNGS QFFITTVPTP HLDGKHVVFG
 Consensus  ekf.denf.. kh...g.lsm anag.ntngs qff..t..t. .ldgkhvvfg

            151                                                        200
CYPH_HUMAN  KVKEGMNIVE AMERFGSRNG KTSKKITIAD CGQLE..... ..........
CYP3_CAEEL  RVVEGLDVVK AVESNGSQSG KPVKDCMIAD CGQLKA.... ..........
CYP4_HUMAN  QVIKGIGVAR ILENVEVKGE KPAKLCVIAE CGELKEGDDG GIFPKDGSGD
CYP4_BOVIN  QVIKGMGVAK ILENVEVKGE KPAKLCVIAE CGELKEGDDW GIFPKDGSGD
PPP5_HUMAN  .......... .......... .......... .......... .....MAMAE
 Consensus  .v..g..... ..e....... k..k...ia. cg.l...... ..........

            201                                                        250
CYPH_HUMAN  .......... .......... .......... .......... ..........
CYP3_CAEEL  .......... .......... .......... .......... ..........
CYP4_HUMAN  SHPDFPEDAD IDLKDVDKIL LITEDLKNIG NTFFKSQNWE MAIKKYAEVL
CYP4_BOVIN  SHPDFPEDAD VDLKDVDKIL LISEDLKNIG NTFFKSQNWE MAIKKYTKVL
PPP5_HUMAN  GERTECAEPP RDEPPADGAL KRAEELKTQA NDYFKAKDYE NAIKFYSQAI
 Consensus  .......... .d....d..l ...e.lk... n..fk....e .aik.y....

            251                                                        300
CYPH_HUMAN  .......... .......... .......... .......... ..........
CYP3_CAEEL  .......... .......... .......... .......... ..........
CYP4_HUMAN  RYVDSSKAVI ETADRAKLQP IALSCVLNIG ACKLKMSNWQ GAIDSCLEAL
CYP4_BOVIN  RYVEGSRAAA EDADGAKLQP VALSCVLNIG ACKLKMSDWQ GAVDSCLEAL
PPP5_HUMAN  .......... ......ELNP SNAIYYGNRS LAYLRTECYG YALGDATRAI
 Consensus  .......... .......l.p .......n.. ...l...... .a......a.

            301                                                        350
CYPH_HUMAN  .......... .......... .......... .......... ..........
CYP3_CAEEL  .......... .......... .......... .......... ..........
CYP4_HUMAN  ELDPSNTKAL YRRAQGWQGL KEYDQALADL KKAQGIAPED KAIQAELLKV
CYP4_BOVIN  EIDPSNTKAL YRRAQGWQGL KEYDQALADL KKAQEIAPED KAIQAELLKV
PPP5_HUMAN  ELDKKYIKGY YRRAASNMAL GKFRAALRDY ETVVKVKPHD KDAKMKYQEC
 Consensus  e.d....k.. yrra.....l .....al.d. .......p.d k.........

            351               370
CYPH_HUMAN  .......... ..........
CYP3_CAEEL  .........: ..........
CYP4_HUMAN  KQKIKAQKDK EKAVYAKMFA
CYP4_BOVIN  KQKIKAQKDK EKAAYAKMFA
PPP5_HUMAN  NKIVK.QKAF ERAIAGDEHK....
 Consensus  ....k.qk.. e.........
```

(*a*)

```
CYP4_BOVIN          LISEDLKNIGNTFFKSQNWEMAIKKYTKVLRYVEGSRAAAEDADGAKLQP
PPP5_HUMAN          KRAEELKTQANDYFKAKDYENAIKFYSQAIELNP..... ......

CYP4_BOVIN          VALSCVLNIGACKLKMSDWQGAVDSCLEALEIDP
PPP5_HUMAN_         SNAIYYGNRSLAYLRTECYGYALGDATRAIELDK

CYP4_BOVIN          SNTKALYRRAQGWQGLKEYDQALADLKKAQEIAPEDKAIQAELLKVKQKIKAQKDKEKAAYAKMFA
PPP5_HUMAN          KYIKGYYRRAASNMALGKFRAALRDYETVVKVKPHDKDAKMKYQECNKIVK QKAFERAIAGDEHK.....
```

(*b*)

Figure 2
(*a*) Sequence alignment of human cyclophilin A (CYPH_HUMAN), nematode cyclophilin 3 (CYP3_CAEEL), human and bovine cyclophilin 40 (CYP4_HUMAN, CYP4_BOVIN) and human protein phosphatase 5 (PPP5_HUMAN). Sequences from the Swiss–Prot database were aligned using the program *MultAlign* and represent the entire sequence of the cyclophilins aligned against the N-terminal 150 amino acids from PP5. The TPRs of PP5 are shown in bold. Numbering applies to the amino-acid sequence of bovine CyP40. (*b*) Alignment of the three TPR motifs of bCyP40 against the N-terminal 150 TPR domain of PP5. The six helices associated with the three TPR motifs determined in the the X-ray structure of PP5 (Das *et al.*, 1998) are underlined. The C-terminal extension of 28 amino acids is the predicted helical calmodulin-binding site. The program *MultAlign* (Corpet, 1988) was used for amino-acid alignment.

3. Discussion

The amino-acid sequence for bCyP40 is shown in Fig. 2(*a*) and shares 94% identity with the human homologue. Mass spectrometry gives a molecular mass of 40500 Da. The calculated molecular mass of the sequence shown in Fig. 2 is 40620 Da, which suggests that the N-terminal methionine residue is removed post-translationally. The calculated molecular mass of the sequence omitting the N-terminal methionine is 40489 Da.

The cyclophilin domain, consisting of the N-terminal 187 residues, is very well conserved between the bovine and human proteins, with differences in only five residues. The N-terminal domain is 61% identical to hCyP18 and shows an even better homology (64%) with the so-called divergent class of cyclophilins (Taylor *et al.*, 1998), which have an additional seven amino-acid loop inserted at positions 60–67 (Fig. 2*a*). Another feature of this divergent class of cyclophilins is that the active-site tryptophan, shown to be important in binding CsA in hCyP18 (Pflugl *et al.*, 1994), is replaced by a histidine. This change may explain the reduction of the IC_{50} values for CsA from 20 n*M* for hCyP18 to 300 n*M* for hCyP40 (Kieffer *et al.*, 1992). The CyP40/CsA/calcineurin K_i was 320 n*M*, compared with the CyP18/CsA/calcineurin K_i of 195 n*M*, and CyP40 was found to compete with CyP18 for binding to calcineurin in the presence of CsA (Kieffer *et al.*, 1993).

The C-terminal 150 amino acids (residues 188–359) of hCyP40, comprising the TPR and putative calmodulin-binding domains, is 31% homologous with the corresponding region of FKBP59. An X-ray structure is available for the TPR domain of the human protein PP5 (Das *et al.*, 1998) which shows that each TPR motif consists of two helices of about 14 amino acids which are linked by a tight turn. The predicted secondary structure of the C-terminal CyP40 domain is shown in Fig. 2(*b*), along with a sequence alignment between human phosphatase and bCyP40. Helical regions of the predicted TPR motifs are also shown.

The PP5 crystal structure (Das *et al.*, 1998) shows that the putative calmodulin-binding domain of CyP40 consists of an α-helix approximately 30 amino acids in length. The secondary-structure prediction (Rost *et al.*, 1994) of the CyP40 sequence gives a strongly predicted C-terminal helix of 28 amino acids in length. This corresponds to a length of 42 Å and is close to the conserved *b* cell dimension of the monoclinic needle form of CyP40. The *b* axis also corresponds to the

long and fast-growing needle axis, and it is reasonable to suggest that it is this calmodulin-binding helix which determines the length of the *b* axis and also the kinetics and morphology of the monoclinic needle crystals.

We thank Violet R. Anderson for the mass spectrometry.

References

Anderson, S. K., Gallinger, S., Roder, J., Frey, J., Young, H. A. & Ortaldo, J. R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 542–546.

Bose, S., Weikl, T., Bugl, H. & Buchner, J. (1996). *Science*, **274**, 1715–1717.

Carrello, A., Ingley, E., Minchin, R. F., Tsai, S. & Ratajczak, T. (1999). *J. Biol. Chem.* **274**, 2682–2689.

Corpet, F. (1988). *Nucleic Acids Res.* **16**, 10881–10890.

Das, A. K., Cohen, P. T. W. & Barford, D. (1998). *EMBO J.* **17**, 1192–1199.

Duina, A. A., Kalton, H. M. & Gaber, R. F. (1998). *J. Biol. Chem.* **273**, 18974–18978.

Duina, A. A., Marsh, J. A. & Gaber, R. F. (1996). *Yeast*, **12**, 943–952.

Freeman, B. C., Toft, D. O. & Morimoto, R. I. (1996). *Science*, **274**, 1718–1720.

Galat, A. & Metcalfe, S. M. (1995). *Prog. Biophys. Mol. Biol.* (1995). **63**, 67–118.

Handschumacher, R. E., Harding, M. W., Rice, J. & Drugge, R. J. (1984). *Science*, **226**, 544–547.

Hoffmann, K. & Handschumacher, R. E. (1995). *Biochem. J.* **307**, 5–8.

Kallen, J., Spitzfalden, C., Zurini, M. G. M., Wider, G., Widmer, H., Wuthrich, K. & Walkinshaw, M. D. (1991). *Nature (London)*, **353**, 276–279.

Kieffer, L. J., Seng, T. W., Li, W., Osterman, D. G., Handschumacher, R. E. & Bayney, R. M. (1993). *J. Biol. Chem.* **268**, 12303–12310.

Kieffer, L. J., Thalhammer, T. & Handschumacher, R. E. (1992). *J. Biol. Chem.* **267**, 5503–5507.

Nair, S. C., Rimerman, R. A., Toran, E. J., Chen, S. Y., Prapapanich, V., Butts, R. N. & Smith, D. F. (1997). *Mol. Cell. Biol.* **17**, 594–603.

Pflugl, G., Kallen, J., Jansonius, J. M. & Walkinshaw, M. D. (1994). *J. Mol. Biol.* **244**, 385–409.

Pflugl, G., Kallen, J., Schirmer, T., Jansonius, J. N., Zurini, M. G. M. & Walkinshaw, M. D. (1993). *Nature (London)*, **361**, 91–94.

Pratt, W. B. & Toft, D. O. (1997). *Endocrine Rev.* **18**, 306–360.

Radanyi, C., Chambraud, B. & Baulieu, E. E. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11197–11201.

Ratajczak, T. & Carrello, A. (1996). *J. Biol. Chem.* **271**, 2961–2965.

Ratajczak, T., Carrello, A., Mark, P. J., Warner, B. J., Simpson, R. J., Moritz, R. L. & House, A. K. (1993). *J. Biol. Chem.* **268,** 13187–13192.

Rost, B., Sander, C. & Schneider, R. (1994). *Comput. Appl. Biosci.* **10**, 53–60.

Taylor, P., Page, A. P., Kondopidis, G., Husi, H. & Walkinshaw, M. D. (1998). *FEBS Lett.* **425**, 261–366.

Weisman, R., Creanor, J. & Fantes, P. (1996). *EMBO J.* **15**, 447–456.

Yokoyama, N., Hayashi, N., Seki, T., Pante, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, Y., Aebi, U., Fukui, M. & Nishimoto, T. (1995). *Nature (London)*, **376**, 184–188.

Young, J. C., Obermann, W. M. J. & Hartl, F. U. (1998). *J. Biol. Chem.* **273**, 18007–18010.