Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A novel interdomain interface in crystallins: structural characterization of the $\beta\gamma$-crystallin from *Geodia cydonium* at 0.99 Å resolution

Alessandro Vergara,[a,b] Marco Grassi,[a] Filomena Sica,[a,b] Elio Pizzo,[c] Giuseppe D'Alessio,[c] Lelio Mazzarella[a,b] and Antonello Merlino[a,b]*

[a]Department of Chemical Sciences, University of Naples 'Federico II', Via Cintia, I-80126 Napoli, Italy, [b]Istituto di Biostrutture e Bioimmagini (CNR), Via Mezzocannone 16, Napoli, Italy, and [c]Dipartimento di Biologia, Università degli Studi di Napoli 'Federico II', Via Cintia, Napoli, Italy

Correspondence e-mail: antonello.merlino@unina.it

Received 20 December 2012
Accepted 4 February 2013

PDB Reference: geodin, 4iau

The $\beta\gamma$-crystallin superfamily includes highly diverse proteins belonging to all of the kingdoms of life. Based on structural topology, these proteins are considered to be evolutionarily related to the long-lived $\beta\gamma$-crystallins that constitute the vertebrate eye lens. This study reports the crystallographic structure at 0.99 Å resolution of the two-domain $\beta\gamma$-crystallin (geodin) from the sponge *Geodia cydonium*. This is the most ancient member of the $\beta\gamma$-crystallin superfamily in metazoans. The X-ray structure shows that the geodin domains adopt the typical $\beta\gamma$-crystallin fold with a paired Greek-key motif, thus confirming the hypothesis that the crystallin-type scaffold used in the evolution of bacteria and moulds was recruited very early in metazoans. As a significant new structural feature, the sponge protein possesses a unique interdomain interface made up by pairing between the second motif of the first domain and the first motif of the second domain. The atomic resolution also allowed a detailed analysis of the calcium-binding site of the protein.

1. Introduction

The $\beta\gamma$-crystallin superfamily comprises diverse proteins that contain the crystallin-type Greek-key β-strand motif (Jaenicke & Slingsby, 2001). β-Crystallins and γ-crystallins from eye lens are the most representative members of this superfamily (Wistow & Piatigorsky, 1988; Lubsen *et al.*, 1988; Wistow, 1990; van Rens *et al.*, 1992; Jaenicke & Slingsby, 2001; Andley, 2007). γ-Crystallins are monomeric (Najmudin *et al.*, 1993), whereas β-crystallins exist as dimers or tetramers (Smith *et al.*, 2007; Fig. 1). The single chain of γ-crystallins and each subunit of the oligomeric β-crystallins are composed of two homologous domains termed the N-domain and C-domain. Each domain is made up of two Greek-key motifs: motifs M1 and M2 in the N-domain and motifs M3 and M4 in the C-domain. The two domains are believed to have evolved by gene duplication and fusion from an ancestral single-domain $\beta\gamma$-crystallin (Lubsen *et al.*, 1988; D'Alessio, 2002).

β-Crystallins and γ-crystallins present highly conserved hydrophobic patches at their interdomain interfaces. In γ-crystallins, the hydrophobic residues Met43, Phe56 and Ile81 in the M2 motif of the N-domain interact with Val132, Leu145 and Val170 in the M4 motif of the C-domain (Fig. 1*e*). Similar interactions occur at the β-crystallin interface between Val43, Val56 and Ile81 and the triad Val132, Leu145 and Ile170 (βB2-crystallin numbering system; Bax *et al.*, 1990). A major role in the oligomerization process of $\beta\gamma$-crystallins is played by the interdomain peptide (Mayr *et al.*, 1994), which adopts a closed conformation in γ-crystallins and an extended conformation in β-crystallins. Other distinctive structural features of

members of the $\beta\gamma$-crystallin superfamily are two characteristics of most β-sheet proteins: the Tyr and Trp corners (Hamill *et al.*, 2000; Aravind *et al.*, 2008). Furthermore, a rather common motif for binding calcium has been found in $\beta\gamma$-crystallins from bacteria, archaea, slime moulds and urochoradates, even though it is absent from most lens β- and γ-crystallins (Mishra *et al.*, 2012).

Although crystallins from eye lens have been extensively studied, to date only a few non-lens $\beta\gamma$-crystallins have been characterized (see, for example, Ray *et al.*, 1997; Aravind, Mishra *et al.*, 2009; Aravind, Suman *et al.*, 2009; Kappé *et al.*, 2010). In particular, the structures of protein S from the bacterium *Myxococcus xanthus* (Bagby *et al.*, 1994), which has two domains, and of spherulin 3a from *Physarum polycephalum* (Kretschmar *et al.*, 1999), *Ciona intestinalis* $\beta\gamma$-crystallin (Ci-$\beta\gamma$-crystallin; Shimeld *et al.*, 2005), nitrollin from *Nitrosospira multiformis* (Aravind, Suman *et al.*, 2009) and ancestral $\beta\gamma$-crystallin from *Methanosarcina acetivorans* (Barnwal *et al.*, 2008), which have a single domain, are known. Recently, single $\beta\gamma$-crystallin domains of proteins from *Clostridium beijerinckii* (clostrillin; Aravind, Mishra *et al.*, 2009), *Flavobacterium johnsoniae* (flavollin; Aravind, Mishra *et al.*, 2009), *Hahella chejeuensis* (Srivastava *et al.*, 2010) and mouse brain (Crybg3; Rajanikanth *et al.*, 2012) have been also reported. Interestingly, when compared with lens $\beta\gamma$-crystallins, some of these non-lens proteins present limited sequence identity and distinct differences in their biophysical properties, such as Ca^{2+}-binding ability (Wenk *et al.*, 1999; Clout *et al.*, 2001; Mayr *et al.*, 1997; Mishra *et al.*, 2012). For these reasons, members of the $\beta\gamma$-crystallin superfamily form an appropriate model system for exploring the limits of the relationship between sequence and structure and its effect on Ca^{2+} binding and stability. Here, we report the atomic resolution (0.99 Å) structure of *Geodia cydonium* $\beta\gamma$-crystallin in its Ca^{2+}-bound form. This protein, termed geodin, has a relatively low (<25% per domain) sequence identity to known $\beta\gamma$-crystallins and is the most ancient member of the $\beta\gamma$-crystallin superfamily (Di Maro *et al.*, 2002). Geodin is a monomeric two-domain protein which is more prone to unfold in the presence of chemical denaturants and high temperatures compared with other $\beta\gamma$-crystallins (Giancola *et al.*, 2005).

The crystal structure, which was solved by single-wavelength anomalous dispersion (Dauter *et al.*, 2002), shows that geodin possesses a unique interdomain interface.

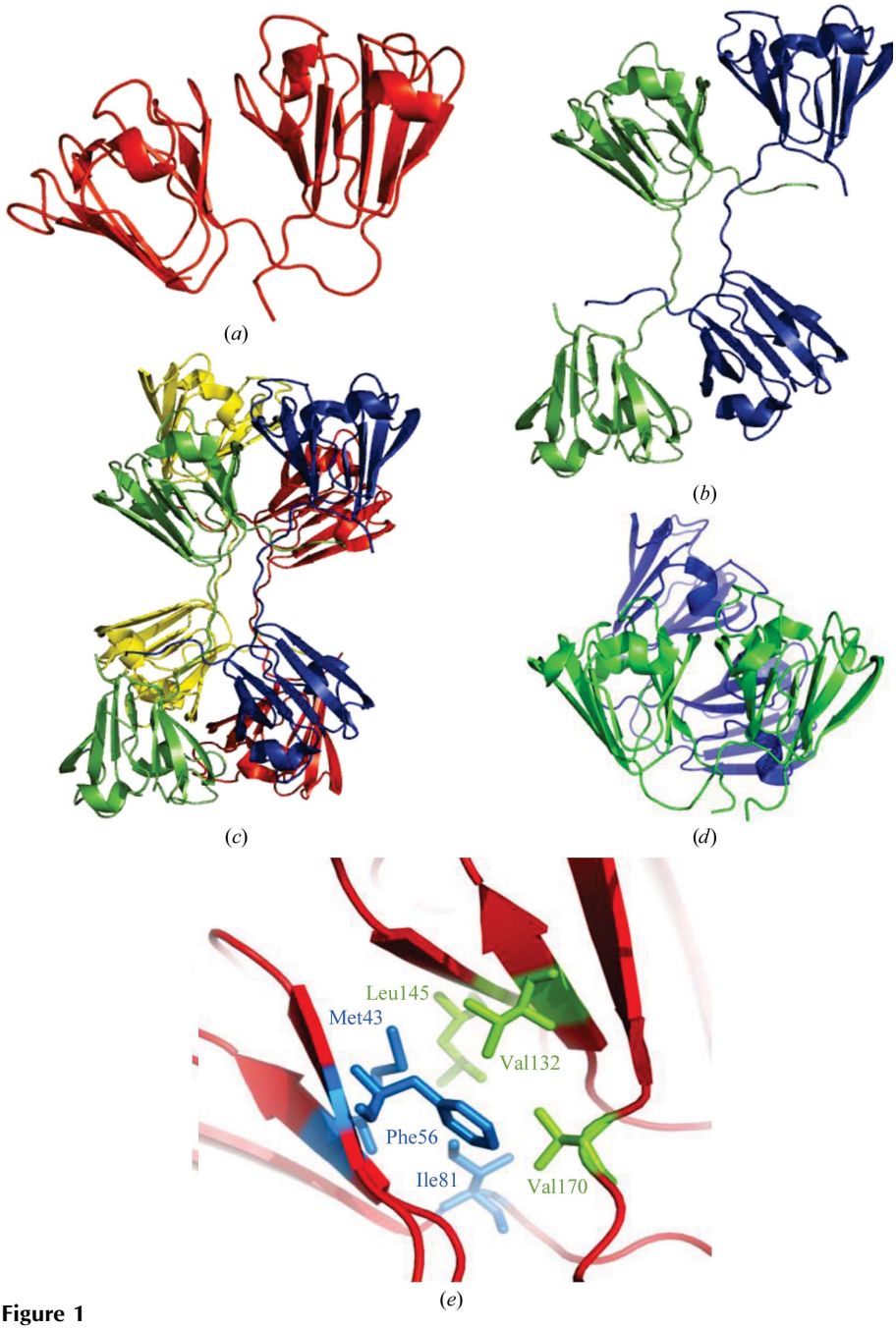

Figure 1
Cartoon representations of the structures of (*a*) bovine γB-crystallin (Najmudin *et al.*, 1993), (*b*) dimeric human βB2-crystallin (Smith *et al.*, 2007), (*c*) tetrameric β-crystallin and (*d*) a β-crystallin showing domain pairing like γ-crystallin (Van Montfort *et al.*, 2003). (*e*) Side chains of residues in the hydrophobic core of the interdomain γ-crystallin interface.

2. Methods

2.1. Crystallization and X-ray diffraction data collection of geodin

Crystals of wild-type geodin and of its selenomethionine derivative (SeMet geodin) were grown in sitting drops at

Table 1
Data-collection and refinement statistics for SeMet geodin.

Values in parentheses are for the highest resolution shell.

Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 24.17, b = 61.30, c = 50.17, β = 95.45
Resolution range (Å)	30.00–0.99 (1.01–0.99)
Unique reflections	80711
Completeness (%)	99.6 (98.4)
$\langle I/\sigma(I)\rangle$	21.0 (3.0)
R_{merge} (%)	6.4 (25.8)
Multiplicity	8.3 (2.1)
Refinement	
Resolution range (Å)	30.00–0.99
Working-set reflections	76488
Test-set reflections	4017
R.m.s.d. from ideality	
Bonds (Å)	0.026
Angles (°)	2.040
R factor (%)	9.8
R_{free}† (%)	12.1
No. of protein atoms	1309
No. of water molecules	370
No. of glycerol molecules	4
No. of Ca^{2+} ions	1
Ramachandran plot, residues in (%)	
Most favoured region (%)	97.5
Additionally allowed region (%)	2.5
Average B factors (Å^2)	
Protein (overall)	8.92
Main-chain atoms	7.58
Side-chain atoms	18.35
Water molecules	29.80
Ca^{2+} ions	10.30
Glycerol atoms	22.55

† Throughout refinement, 5% of the total reflections were kept aside for the calculation of R_{free}.

277 K by mixing 2 µl protein solution at 2.4 mg ml^{-1} with 2 µl reservoir solution consisting of 0.2 *M* ammonium sulfate, 17%(*w*/*v*) PEG 8000, 2 m*M* dithiothreitol, 50 m*M* Tris–HCl pH 7.6 (Vergara *et al.*, 2008).

X-ray diffraction data were collected from a single crystal of SeMet geodin to 0.99 Å resolution on beamline XRD1 at the Elettra synchrotron, Trieste, Italy. The crystals belonged to the monoclinic space group $P2_1$, with unit-cell parameters a = 24.17, b = 61.30, c = 50.175 Å, β = 95.45°. Data-collection statistics for the crystals are given in Table 1. Using a conventional source, X-ray diffraction data were also collected for wild-type geodin crystals at room temperature (300 K). These crystals were isomorphous to those of SeMet geodin, with unit-cell parameters a = 27.98, b = 61.84, c = 50.57 Å, β = 97.21°, and diffracted to 1.9 Å resolution.

2.2. Structure determination and refinement

The structure of SeMet geodin was solved by the SAD method using the anomalous scattering of the Se atoms. The positions of the Se atoms were determined by *SHELXD* after analyzing the structure factors using *SHELXC*. Phase calculation and density modification were then performed using *SHELXE*. Initial phasing trials applying 20 cycles of density modification led to an interpretable and very well defined experimental electron-density map that enabled automated model building of 147 of the 163 residues using *ARP*/*wARP* (Langer *et al.*, 2008). The partial model was manually completed using *O* (Jones *et al.*, 1991) alternated with structure refinement by restrained conjugated-gradient least-squares (CGLS) minimization in *SHELXL*-97 (Sheldrick, 2008). A randomly selected 5% of the reflections were used for cross-validation. After initial isotropic refinement, anisotropic atomic displacement parameters were refined. Inspection of the OMIT $F_o - F_c$ electron-density maps allowed the addition of water, glycerol molecules and a calcium ion and the building of alternate conformations for 13 of the 163 side chains. The final refinement included all H atoms in calculated positions. Standard restraints were applied to the geometrical and isotropic or anisotropic displacement parameters. The R factor and R_{free} converged to final values of 9.8 and 12.1%, respectively, for data with $F_o > 4\sigma(F_o)$ and of 10.6 and 12.8%, respectively, for all data. The structure of the wild-type protein at room temperature was solved by molecular replacement with *AMoRe* (Navaza, 1994) using the coordinates of the refined structure of SeMet geodin as the search probe and was refined with *CNS* (Brunger, 2007). The final values of the R factor and R_{free} converged to 19.8 and 22.1%, respectively, for data with $F_o > 2\sigma(F_o)$. The stereochemical quality of the structures was checked and validated using *PROCHECK* (Laskowski *et al.*, 1993). The figures were generated using *PyMOL* (http://www.pymol.org). Atomic coordinates and experimental structure factors for SeMet geodin have been deposited in the Protein Data Bank with accession code 4iau.

3. Results and discussion

3.1. Model quality

The crystal structure of SeMet geodin was determined by the SAD method (Dauter *et al.*, 2002) using the anomalous scattering of the Se atoms and was refined to 0.99 Å resolution. The final model contained 161 residues (Ser2–Glu162), a calcium ion, four glycerol molecules and 370 fully occupied water molecules. All nonglycine residues lie in the most favoured (97.5%) or additionally allowed (2.5%) regions of the Ramachandran plot. Statistics of data collection and

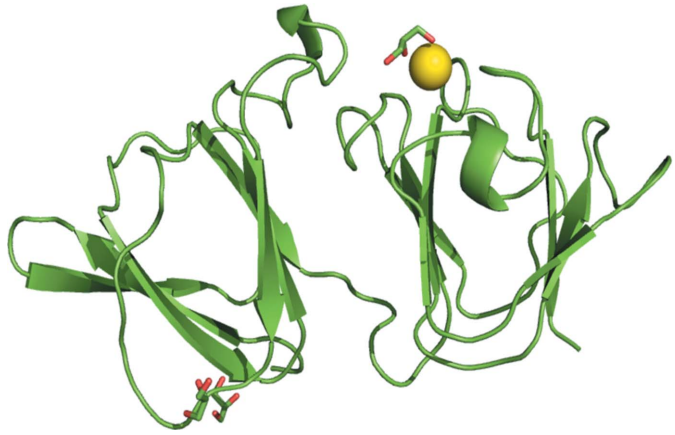

Figure 2
Cartoon representation of the geodin structure. The calcium ion is shown as a yellow sphere. Glycerol molecules are also shown.

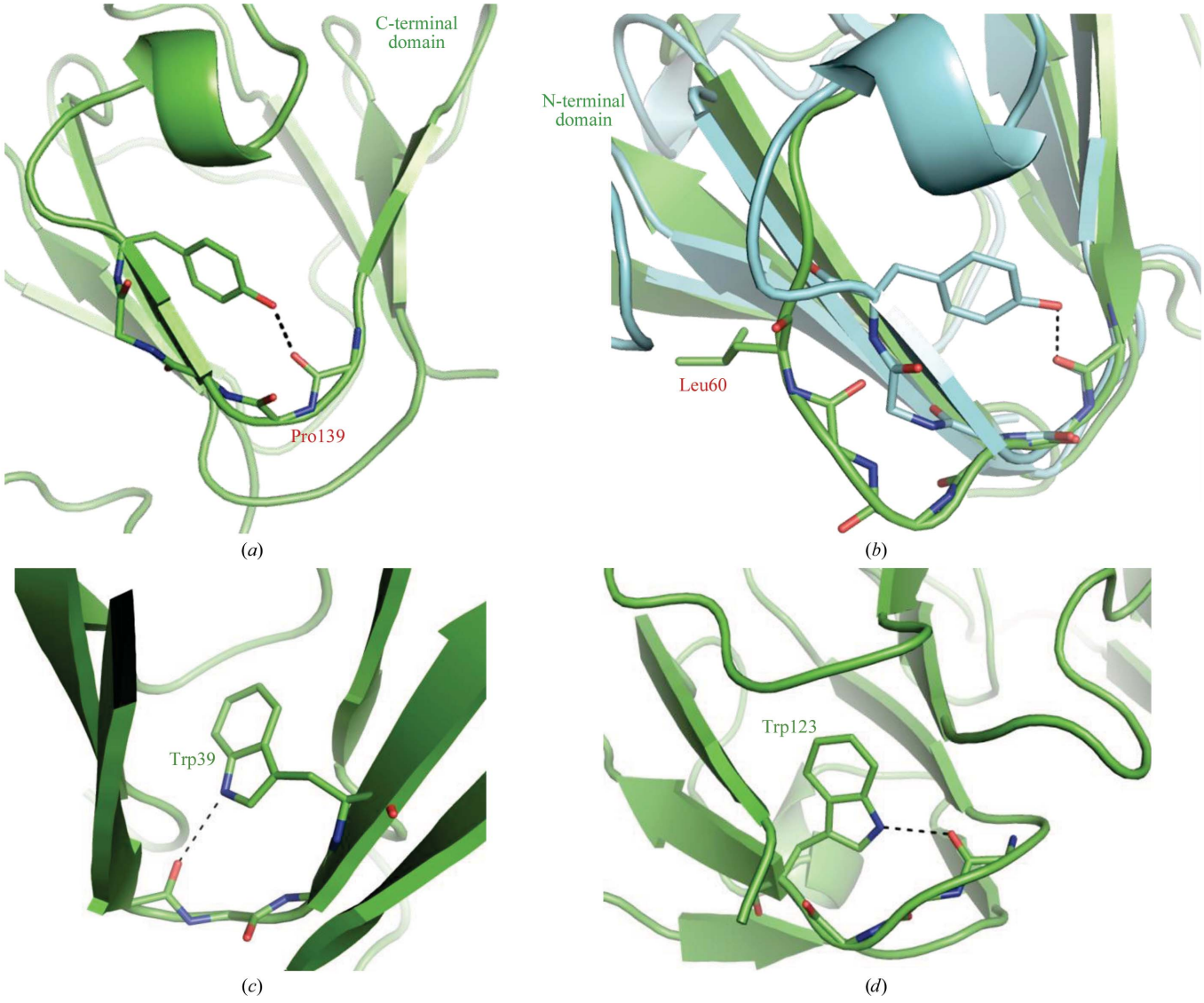

Figure 3
(*a*) The Tyr corner in the C-domain of geodin. (*b*) Superimposition of the N-domain of bovine γB-crystallin (cyan) and that of geodin (green). In the latter, Tyr is replaced by Leu60 and the Tyr corner is missing. (*c*, *d*) The Trp corner in the N- and C-domains of geodin. In the N-domain, the indole NH of Trp39 forms a hydrogen bond to the main-chain carbonyl of Lys36; in the C-domain, the indole NH of Trp123 forms a hydrogen bond to the main-chain carbonyl of Glu120. For sake of clarity side-chain atoms of Lys36 and Glu120 have been omitted.

processing and relevant refinement parameters are presented in Table 1.

The electron-density maps are very well defined, except for residues SeMet38 and for the terminal residues SeMet1 and Gln163. Holes are clearly observed in the Phe, Tyr, Trp and Pro rings (Supplementary Fig. S1[1]). The high definition of the electron-density maps allowed us to discriminate between the N and C atoms of His residues and between the N and O atoms of Asn and Gln side chains. The atomic resolution analysis led us to introduce disordered conformations and H atoms into a fully anisotropic model. Alternate conformations were modelled for 13 residues, namely Ser2, Thr3, Lys5, Ser14, Thr26, SeMet38, Phe41, Lys59, Ser64, SeMet67, Lys80, Lys97 and Thr136. Inspection of the positions occupied by solvent allowed the detection of rings formed by three, four, five, six and seven water molecules (Supplementary Fig. S2), as observed in other ultrahigh-resolution structures of proteins (Esposito *et al.*, 2000).

3.2. Overall fold and structural homology

The structure of SeMet geodin is constituted by four Greek-key motifs, each of approximately 40 residues, arranged in two domains: the N-domain and the C-domain (Fig. 2). The structures of both the N-domain and the C-domain are composed of seven strands connected by loops and turns and by two short helices. In the overall topology, the β-strands are separated into two sheets and these sheets form the Greek-key motif. In the N-domain, the β-strands are termed *A* (residues 5–9), *B* (residues 15–18), *C* (residues 31–26), *D* (residues 38–

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: LV5035). Services for accessing this material are described at the back of the journal.

42), *E* (residues 57–61), *F* (residues 67–68) and *G* (residues 75–78). In the C-domain, the β-strands are termed *A*′ (residues 86–89), *B*′ (residues 98–101), *C*′ (residues 115–118), *D*′ (residues 123–126), *E*′ (residues 135–138), *F*′ (residues 141–143) and *G*′ (residues 158–161). In both domains there are folded hairpin turns between strands *A* and *B* and between strands *D* and *E*. The two domains face each other in the global topology. Hydrophobic and aromatic side chains, including SeMet38, Ile40, Tyr42, Trp57, Val79, Phe82, Tyr89, His91, Phe94 and Leu117, are packed in the large interface area of the β-sheets. Eight hydrogen bonds are found at the inter-domain interface (see below).

As determined by the *DALI* server (Holm *et al.*, 2008), the N-domain shows highest structural similarity to human βB2-crystallin [PDB entry 1ytq; C^{α} root-mean-square deviation (r.m.s.d.) 1.6 Å; sequence identity of 21% over 70 residues; Smith *et al.*, 2007], whereas the C-domain has highest structural similarity to human γD-crystallin (PDB entry 1hk0; r.m.s.d. 1.2 Å; sequence identity of 22% over 79 residues; Basak *et al.*, 2003) (Supplementary Table S1).

3.3. Tyr and Trp corners

The tyrosine corner is a conserved structural feature of most β-sheet proteins which is conventionally present in the β-motif of the βγ-crystallins (Hamill *et al.*, 2000; Aravind *et al.*, 2008). In this feature, a tyrosine hydroxyl makes a hydrogen bond to the main-chain carbonyl four residues ahead and may act as a nucleus during the folding process of the protein. Analysis of the crystal structure of SeMet geodin reveals that the Tyr corner is conserved in the C-domain but not in the N-domain (Figs. 3*a* and 3*b*). When the structure and the sequence of the geodin N-domain are aligned with those of other βγ-crystallins a deletion is observed in the region corresponding to the Tyr corner and the Tyr is replaced by a Leu (Leu60). Although the Tyr corner is missing, the N-domain folds into the typical β-sandwich, showing that this feature may not be essential for the βγ-crystallin fold. The lack of the Tyr corner is also associated with a Tyr→Leu replacement in the case of the βγ-crystallin domains of both protein absent in melanoma-1 (AIM1; PDB entry 3cw3; Aravind *et al.*, 2008) and the crystallin from *Yersinia pestis* (Jobby & Sharma, 2005).

Another structural feature of βγ-crystallins is the Trp corner (Hamill *et al.*, 2000). This is an evolutionarily conserved motif which has been related to protein function in the vertebrate lens crystallins, since Trp mutations seem to be implicated in many cases of congenital cataracts (Talla *et al.*, 2006, 2008). In this motif, the indole NH forms a hydrogen bond to a main-chain carbonyl and can form strong hydrophobic interactions in the domain core. This structural feature is indeed retained in both the N- and C-domains (Figs. 3*c* and 3*d*).

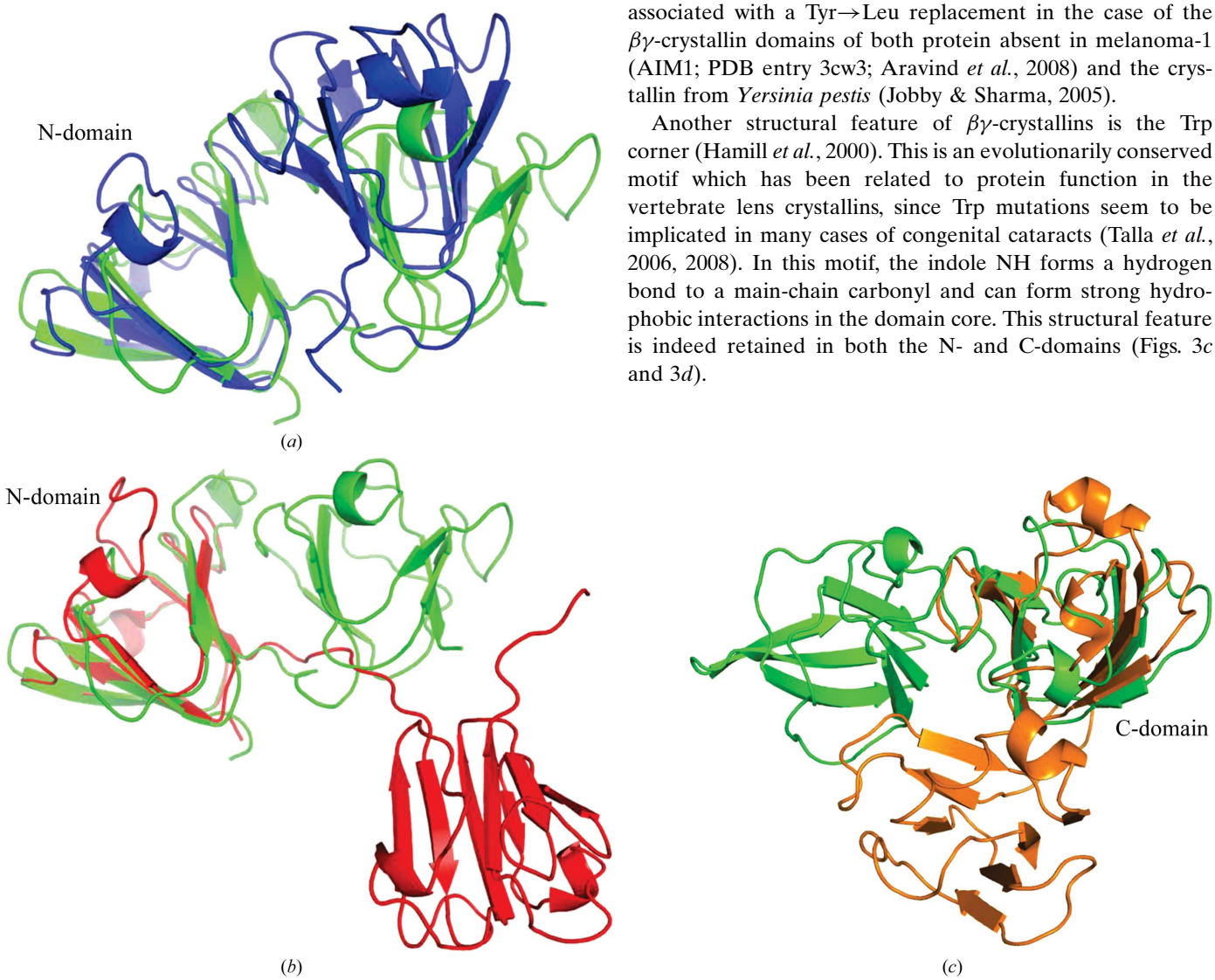

Figure 4
Structural superimposition of geodin (green) and other βγ-crystallins: (*a*) bovine γB-crystallin (Najmudin *et al.*, 1993; blue), (*b*) human βB2-crystallin (Smith *et al.*, 2007; red) and (*c*) protein S (Bagby *et al.*, 1994; orange).

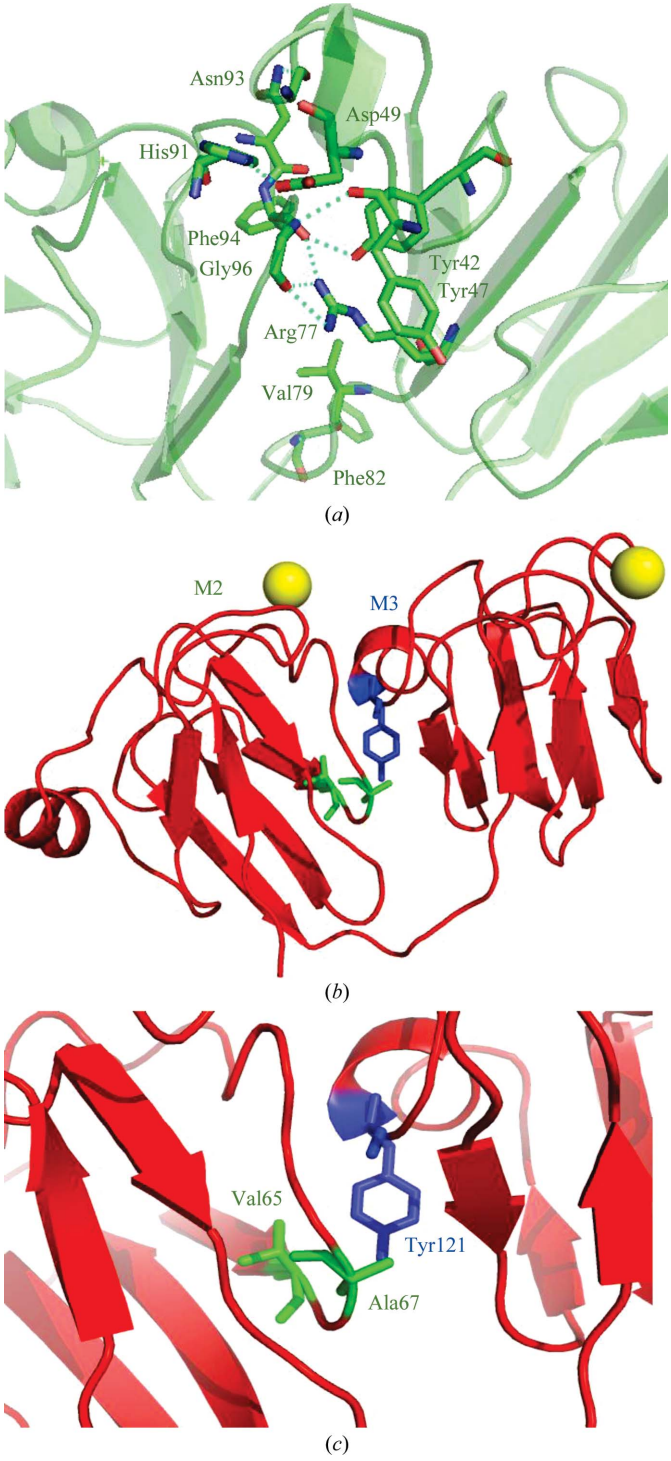

Figure 5
(*a*) Hydrogen bonds at the geodin interdomain interface. (*b*) Cartoon representation of protein S. The calcium ion is shown as a yellow sphere and interface residues are highlighted. (*c*) Side chains of residues at the hydrophobic core of the protein S interdomain interface.

Table 2
Interdomain hydrogen bonds and hydrophobic interactions in geodin.

N-domain	C-domain	Distance (Å)
Hydrogen bonds		
Tyr42 OH	Phe94 O	2.64
Tyr47 O	Gly96 N	2.86
Asp49 $O^{\delta1}$	His91 $N^{\varepsilon2}$	2.80
Gly53 O	Asn93 $N^{\delta2}$	3.16
Arg77 $N^{\eta2}$	Gly96 O	2.75
Arg77 $N^{\eta2}$	Phe94 O	2.88
Arg77 $N^{\eta1}$	Gly96 O	2.96
Val79 O	Phe82 N	3.07
Hydrophobic interactions		
Gly37 C^{α}	Phe82 $C^{\delta2}$	3.72
Gly37 C^{α}	Phe82 $C^{\varepsilon2}$	3.74
SeMet38 C^{γ} A	Phe82 $C^{\delta1}$	3.66
SeMet38 C^{γ} A	Phe82 $C^{\varepsilon1}$	3.90
SeMet38 C^{ε} B	Phe82 $C^{\delta1}$	3.30
Ile40 $C^{\delta1}$	Leu117 $C^{\delta1}$	3.70
Ile40 $C^{\delta1}$	Phe94 C^{β}	3.63
Tyr42 $C^{\varepsilon2}$	Phe94 C	3.87
Gly52 C^{α}	His91 C^{γ}	3.65
Gly52 C^{α}	His91 $C^{\delta2}$	3.74
Gly52 C^{α}	His91 $C^{\varepsilon1}$	3.89
Gly52 C	His91 C^{β}	3.51
Gly52 C	His91 C^{γ}	3.49
Gly52 C	His91 $C^{\delta2}$	3.85
Gly53 C^{α}	His91 C^{β}	3.87
Trp57 $C^{\varepsilon2}$	Asn93 C^{β}	3.65
Trp57 $C^{\varepsilon2}$	Asn93 C^{γ}	3.75
Trp57 $C^{\delta1}$	Asn93 C^{β}	3.60
Trp57 $C^{\zeta3}$	Phe94 $C^{\delta2}$	3.75
Trp57 $C^{\zeta3}$	Phe94 $C^{\varepsilon2}$	3.63
Trp57 $C^{\eta2}$	Phe94 $C^{\varepsilon2}$	3.82
Lys59 C^{δ} A	Leu117 $C^{\delta1}$	3.85
Lys59 C^{ε} B	Leu117 $C^{\delta2}$	3.90
Arg77 C^{ζ}	Tyr89 $C^{\varepsilon1}$	3.87
Val79 $C^{\gamma2}$	Leu117 $C^{\delta1}$	3.70

3.4. Domain pairing in geodin and in the $\beta\gamma$-crystallin family

As in γ-crystallins and protein S, geodin has two domains and is monomeric. γ-Crystallins have a symmetric interdomain interface made up by the second motifs of both domains, whereas protein S has a different interface made up by the second motif of the first domain and the first motif of the second domain.

The succession of motifs at the interdomain SeMet geodin interface is similar to that observed in protein S and differs from those found in vertebrate $\beta\gamma$-crystallins (Wistow *et al.*, 1983). However, the orientation of the domains with respect to each other is not the same (Fig. 4). A detailed comparison of the features of the interdomain interface in SeMet geodin and representatives of the different crystallins is reported in Supplementary Table S2. The SeMet geodin domain interface is larger than that of protein S (about 520 Å^2), with a surface area of about 680 Å^2, whereas it is smaller than those of vertebrate γ-crystallins (average value of about 800 Å^2). The comparison also reveals that the SeMet geodin interdomain interface is stabilized by a greater number of hydrogen bonds and hydrophobic interactions compared with the other crystallins (Table 2 and Fig. 5*a*). The interface interactions that are conserved in the known γ-crystallins (Fig. 1*e*) are not observed in SeMet geodin. This observation is not surprising since some of the hydrophobic residues involved in the formation of the γ-crystallin interface (Val132, Leu145 and Val170 of motif M4) are replaced by nonhydrophobic residues in geodin (Arg124, Arg137 and Glu162). Similarly, the interface residues found in protein S (Val65, Ala67 and Tyr121; Fig. 5*b*) are nonconserved in geodin (Gly54, Trp57 and Pro106). On the basis of these

results, it can be surmised that the smaller buried surface at the interdomain interface of SeMet geodin could play a role in its decreased stability to thermal and denaturing agents compared with other $\beta\gamma$-crystallins (Giancola *et al.*, 2005). This finding supports the hypothesis that the interdomain surface is very important for maintaining the long lifetime of the proteins in the lens nucleus (Das *et al.*, 2010).

3.5. Calcium-binding site

The ability of geodin to bind calcium ion has been demonstrated using calorimetric and spectroscopic measurements (Giancola *et al.*, 2005). The atomic resolution structure of SeMet geodin allows the clear identification of the calcium-binding site. It is located on the protein surface at a position topologically equivalent to that reported for Ci-$\beta\gamma$-crystallin (Shimeld *et al.*, 2005) and for other nonvertebrate $\beta\gamma$-crystallins (Bagby *et al.*, 1994; Rajini *et al.*, 2001; Jobby & Sharma, 2005, 2007). However, the pattern of calcium coordination in SeMet geodin is slightly different compared with those observed in other $\beta\gamma$-crystallins (Mishra *et al.*, 2012; Supplementary Table S3). In particular, the calcium ion is hepta-coordinated and is complexed in a ligand field made up by the O atoms of the Ser114 and Asp154 side chains, the backbone carbonyl O atoms of Gly112 and Lys90, one water molecule and two O atoms of a glycerol molecule (Fig. 6*a*). Interestingly, the Ca^{2+}-coordination geometry is perfectly conserved in the structure of wild-type protein refined from data collected at room temperature, with two water molecules occupying the positions of the two glycerol O atoms (Fig. 6*b*).

3.6. Domain-connecting peptide

The peptide that links the two domains of $\beta\gamma$-crystallins (the interdomain peptide or domain-connecting peptide) plays a crucial role in their association (Mayr *et al.*, 1994) and hence in the formation of oligomeric species. It has been suggested that oligomeric β-crystallins have a proline at position 80, which precedes the connecting interdomain peptide in the sequence (Trinkl *et al.*, 1994). This proline forces the interdomain peptide to adopt an extended conformation, which in turn favours oligomerization (Bax *et al.*, 1990; Trinkl *et al.*, 1994). In contrast, in γ-crystallins, where the domain-connecting peptide is shorter and adopts a closed conformation, intramolecular interdomain association is favoured (Sharma *et al.*, 1990). In γ-crystallins, most of the residues between position 80 and 87 are in the polyproline conformation, with a Gly residue (Gly86) allowing a sharp turn.

Multiple sequence alignment shows that geodin has a proline (Pro78) corresponding to Pro80 of the β-crystallins (Supplementary Table S4). Therefore, the determination of the crystal structure of SeMet geodin, which is a monomer, provides an interesting case to probe the existence of a relationship between the presence of proline and the quaternary structure adopted by crystallins. It has been suggested that geodin has a short inter-domain linker peptide (only three residues; D'Alessio, 2002). However, structural determination reveals that in geodin, as in the other $\beta\gamma$-crystallins (Blundell *et al.*, 1981; Wistow *et al.*, 1983; D'Alessio, 2002), the interdomain peptide is constituted of five residues (Lys80–Val84). A comparison between the structure of the geodin domain-connecting peptide and that of the same region in other members of the $\beta\gamma$-crystallin superfamily is shown in Supplementary Fig. S3. In SeMet geodin, the tetrapeptide Thr81–Val84 adopts a type II turn conformation that forces the two domains into a compact structure. This interesting result undermines the hypothesis of the role of Pro80 in the formation of the dimeric structure of β-crystallins, in agreement with a previous observation (Van Montfort *et al.*, 2003).

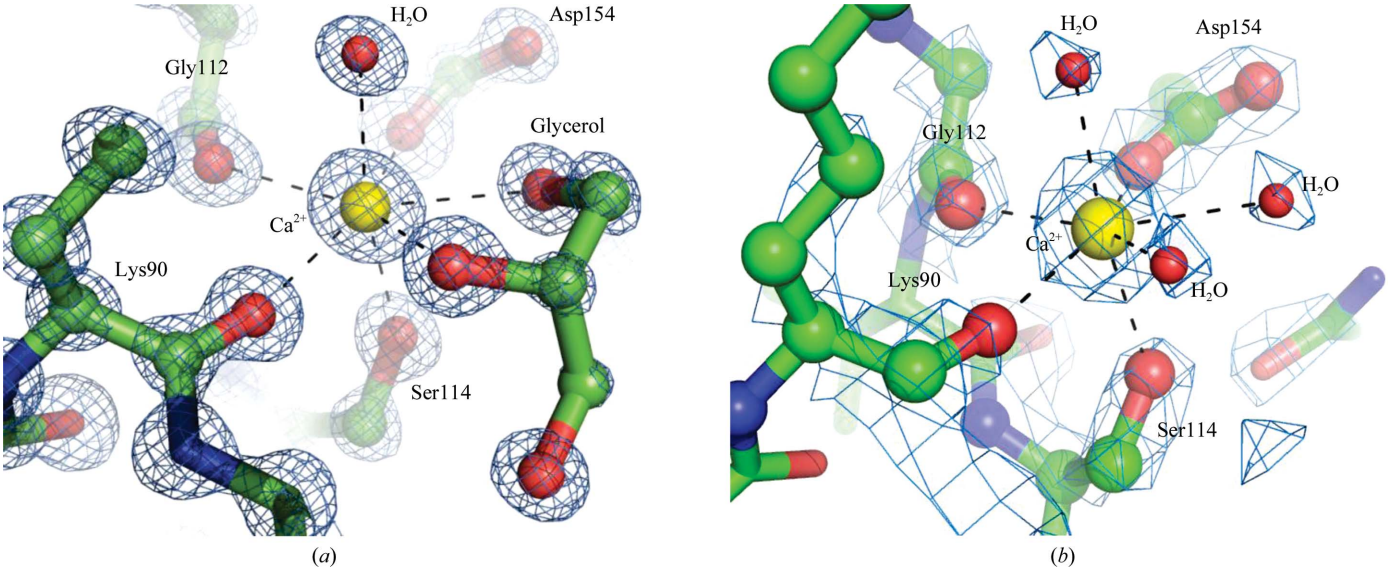

Figure 6
Calcium-binding site in the X-ray structure of SeMet geodin solved at 0.99 Å resolution using data collected at 100 K (*a*) and in the structure of wild-type geodin solved at 1.90 Å resolution using data collected at 300 K (*b*). The $2F_o - F_c$ electron-density map is contoured at 1σ.

4. Conclusions

The atomic resolution structure of geodin, a protein identified in the primitive organism *G. cydonium*, has been determined. The structure retains the characteristic crystallin fold, despite this geodin has a very low sequence identity to the other members of the $\beta\gamma$-crystallin superfamily. This finding supports the hypothesis that this fold can be adopted by proteins with limited sequence identity (Aravind *et al.*, 2008). As a significantly new structural feature, geodin presents an interdomain interface that has never been observed, with association of the M2 and M3 motifs. Such an arrangement could be linked to the functional role of geodin, which should be distinct from that of the $\beta\gamma$-crystallins of the eye lens and is currently unknown.

In geodin, the sequence preceding the domain-connecting peptide shares the proline at position 80 with β-crystallins. However, geodin is monomeric, as are the γ-crystallins. These results underscore the importance of the compactness of the γ-crystallin linker and the role of the interdomain interface in defining the dimeric structure of β-crystallins.

We thank Giosuè Sorrentino and Maurizio Amendola for technical assistance and Elettra Trieste for providing synchrotron-radiation facilities and travel grants to AM and AV.

References

Andley, U. P. (2007). *Prog. Retin. Eye Res.* **26**, 78–98.
Aravind, P., Mishra, A., Suman, S. K., Jobby, M. K., Sankaranarayanan, R. & Sharma, Y. (2009). *Biochemistry*, **48**, 12180–12190.
Aravind, P., Suman, S. K., Mishra, A., Sharma, Y. & Sankaranarayanan, R. (2009). *J. Mol. Biol.* **385**, 163–177.
Aravind, P., Wistow, G., Sharma, Y. & Sankaranarayanan, R. (2008). *J. Mol. Biol.* **381**, 509–518.
Bagby, S., Harvey, T. S., Eagle, S. G., Inouye, S. & Ikura, M. (1994). *Structure*, **2**, 107–122.
Barnwal, R. P., Jobby, M. K., Devi, K. M., Sharma, Y. & Chary, K. V. (2008). *J. Mol. Biol.* **386**, 675–689.
Basak, A., Bateman, O., Slingsby, C., Pande, A., Asherie, N., Ogun, O., Benedek, G. B. & Pande, J. (2003). *J. Mol. Biol.* **328**, 1137–1147.
Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L. & Slingsby, C. (1990). *Nature (London)*, **347**, 776–780.
Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. & Wistow, G. (1981). *Nature (London)*, **289**, 771–777.
Brunger, A. T. (2007). *Nature Protoc.* **2**, 2728–2733.
Clout, N. J., Kretschmar, M., Jaenicke, R. & Slingsby, C. (2001). *Structure*, **9**, 115–124.
D'Alessio, G. (2002). *Eur. J. Biochem.* **269**, 3122–3130.
Das, P., King, J. A. & Zhou, R. (2010). *Protein Sci.* **19**, 131–140.
Dauter, Z., Dauter, M. & Dodson, E. J. (2002). *Acta Cryst.* D**58**, 494–506.
Di Maro, A., Pizzo, E., Cubellis, M. V. & D'Alessio, G. (2002). *Gene*, **299**, 79–82.
Esposito, L., Vitagliano, L., Sica, F., Sorrentino, G., Zagari, A. & Mazzarella, L. (2000). *J. Mol. Biol.* **297**, 713–732.
Giancola, C., Pizzo, E., Di Maro, A., Cubellis, M. V. & D'Alessio, G. (2005). *FEBS J.* **272**, 1023–1035.
Hamill, S. J., Cota, E., Chothia, C. & Clarke, J. (2000). *J. Mol. Biol.* **295**, 641–649.
Holm, L., Kääriäinen, S., Rosenström, P. & Schenkel, A. (2008). *Bioinformatics*, **24**, 2780–2781.
Jaenicke, R. & Slingsby, C. (2001). *Crit. Rev. Biochem. Mol. Biol.* **36**, 435–499.
Jobby, M. K. & Sharma, Y. (2005). *J. Biol. Chem.* **280**, 1209–1216.
Jobby, M. K. & Sharma, Y. (2007). *FEBS J.* **274**, 4135–4147.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kappé, G., Purkiss, A. G., van Genesen, S. T., Slingsby, C. & Lubsen, N. H. (2010). *J. Mol. Evol.* **71**, 219–230.
Kretschmar, M., Mayr, E. M. & Jaenicke, R. (1999). *Biol. Chem.* **380**, 89–94.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lubsen, N. H., Aarts, H. J. & Schoenmakers, J. G. (1988). *Prog. Biophys. Mol. Biol.* **51**, 47–76.
Mayr, E. M., Jaenicke, R. & Glockshuber, R. (1994). *J. Mol. Biol.* **235**, 84–88.
Mayr, E. M., Jaenicke, R. & Glockshuber, R. (1997). *J. Mol. Biol.* **269**, 260–269.
Mishra, A., Suman, S. K., Srivastava, S. S., Sankaranarayanan, R. & Sharma, Y. (2012). *J. Mol. Biol.* **415**, 75–91.
Najmudin, S., Nalini, V., Driessen, H. P. C., Slingsby, C., Blundell, T. L., Moss, D. S. & Lindley, P. F. (1993). *Acta Cryst.* D**49**, 223–233.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Rajanikanth, V., Srivastava, S. S., Singh, A. K., Rajyalakshmi, M., Chandra, K., Aravind, P., Sankaranarayanan, R. & Sharma, Y. (2012). *Biochemistry*, **51**, 8502–8513.
Rajini, B., Shridas, P., Sundari, C. S., Muralidhar, D., Chandani, S., Thomas, F. & Sharma, Y. (2001). *J. Biol. Chem.* **276**, 38464–38471.
Ray, M. E., Wistow, G., Su, Y. A., Meltzer, P. S. & Trent, J. M. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 3229–3234.
Rens, G. L. van, de Jong, W. W. & Bloemendal, H. (1992). *Mol. Biol. Rep.* **16**, 1–10.
Sharma, A. K., Minke-Gogl, V., Gohl, P., Siebendritt, R., Jaenicke, R. & Rudolph, R. (1990). *Eur. J. Biochem.* **194**, 603–609.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Shimeld, S. M., Purkiss, A. G., Dirks, R. P., Bateman, O. A., Slingsby, C. & Lubsen, N. H. (2005). *Curr. Biol.* **15**, 1684–1689.
Smith, M. A., Bateman, O. A., Jaenicke, R. & Slingsby, C. (2007). *Protein Sci.* **16**, 615–625.
Srivastava, A. K., Sharma, Y. & Chary, K. V. (2010). *Biochemistry*, **49**, 9746–9755.
Talla, V., Narayanan, C., Srinivasan, N. & Balasubramanian, D. (2006). *Invest. Ophthalmol. Vis. Sci.* **47**, 5212–5217.
Talla, V., Srinivasan, N. & Balasubramanian, D. (2008). *Invest. Ophthalmol. Vis. Sci.* **49**, 3483–3490.
Trinkl, S., Glockshuber, R. & Jaenicke, R. (1994). *Protein Sci.* **3**, 1392–1400.
Van Montfort, R. L., Bateman, O. A., Lubsen, N. H. & Slingsby, C. (2003). *Protein Sci.* **12**, 2606–2612.
Vergara, A., Merlino, A., Pizzo, E., D'Alessio, G. & Mazzarella, L. (2008). *Acta Cryst.* D**64**, 167–171.
Wenk, M., Baumgartner, R., Holak, T. A., Huber, R., Jaenicke, R. & Mayr, E. M. (1999). *J. Mol. Biol.* **286**, 1533–1545.
Wistow, G. (1990). *J. Mol. Evol.* **30**, 140–145.
Wistow, G. J. & Piatigorsky, J. (1988). *Annu. Rev. Biochem.* **57**, 479–504.
Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P. & Blundell, T. (1983). *J. Mol. Biol.* **170**, 175–202.