Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A comparison of refined X-ray structures of hydrogenated and perdeuterated rat γE-crystallin in H_2O and D_2O

Jean-Baptiste Artero,[a] Michael Härtlein,[a] Sean McSweeney[b] and Peter Timmins[a]*

[a]Institut Laue Langevin, 6 Rue Jules Horowitz, 38042 Grenoble CEDEX 9, France, and [b]European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, 38042 Grenoble CEDEX 9, France

Correspondence e-mail: timmins@ill.fr

Received 12 May 2005
Accepted 11 September 2005

PDB References: hydrogenated γE crystallin in H_2O solvent, 1zgt, r1zgtsf; hydrogenated γE crystallin in D_2O solvent, 1zie, r1ziesf; deuterated γE crystallin in H_2O solvent, 1zir, r1zirsf; deuterated γE crystallin in D_2O solvent, 1ziq, r1ziqsf.

Rat γE-crystallin was overexpressed, purified under different labelling conditions and crystallized and X-ray data were collected at resolutions between 1.71 and 1.36 Å. The structures were determined by molecular replacement. In these structures, the *cd* loop of the Greek-key motif 3, which is the major structural key motif of the two phase-transition groups of γ-crystallins, presents a double conformation. The influence of the perdeuteration on the protein structure was determined by comparison of the atomic positions and temperature factors of the different models. The perdeuterated proteins have a similar structure to their hydrogenated counterparts, but partial or full deuteration may have some effect on the atomic *B*-factor values.

1. Introduction

All vertebrate eye lenses are transparent and have a refractive power that depends on their having a smooth refractive-index gradient for visible light. This is achieved by the regular arrangement of the fibre cell and by the differential expression of lens-specific proteins, the crystallins, and by their organization and their interactions with membranes and cytoskeleton (Benedek, 1971; Delaye & Tardieu, 1983). A reversible opacification of young lenses, obtained on cooling and rewarming and named 'cold cataract', has been observed (Benedek *et al.*, 1979; Ishimoto *et al.*, 1979). A remarkable resemblance between the appearance of opacity in lysozyme–salt water mixtures and the development of opacity in cold cataract in the young rat lens has been postulated to be strong evidence that 'cold cataract' is fundamentally a phase separation of the protein–water binary mixture in the lens (Tanaka *et al.*, 1977). This temperature-dependent behaviour was observed in the opacification of several different lenses such as Philly mouse cataract (Clark & Carper, 1987) or galactosaemic cataract in young rats (Ishimoto *et al.*, 1979). Previous work suggested that this phase separation was driven by the attractive forces between γ-crystallin molecules in solution (Benedek *et al.*, 1979; Tardieu *et al.*, 1992) and certain members of this family show this phenomenon near body temperature (Siezen *et al.*, 1985). With recent advances in the structure determination of several γ-crystallins (Chirgadze *et al.*, 1996; Kumaraswamy *et al.*, 1996; Norledge *et al.*, 1997; Basak *et al.*, 2003), the molecular interactions of γ-crystallins remains to be elucidated in order to understand how these proteins phase-separate at close to body temperature and can thereby cause light scattering in certain cataracts.

Although there is high sequence identity and structural similarity between the different γ-crystallins, two distinct classes can be characterized according to the temperature of their phase separation at a high concentration (250 mg ml^{-1}).

Rat γABC-, human γC- and bovine γB-γIII$_b$-crystallins were characterized as proteins having a low phase-separation temperature, T_c, of ~278 K (Thomson *et al.*, 1987; Broide *et al.*, 1991), whereas rat γEF-, human γD- and bovine γIV$_a$-crystallins were classified as high-T_c proteins ($T_c \simeq 328$ K; Siezen *et al.*, 1988). At temperatures well below T_c, the protein concentration in the high-density phase approaches that found in the centre of the eye lens (Philipson, 1969) and also in protein crystals (Broide *et al.*, 1991; Berland *et al.*, 1992). Therefore protein–protein, water–water and protein–water interactions may be deduced from interactions in the crystal lattice.

All γ-crystallin structures obtained to date possess numerous surface ion pairs that lead to a high surface-charge density but low overall charge as the positive and negative charges are well balanced. Arginines frequently form chains of interacting charges with aspartates in a manner that may provide extra stability to the protein structure. On the other hand, an increased number of charged groups might be expected to interact with water molecules in a more effective way than non-charged groups; thus, their organization as a series of ion pairs may decrease this ordering effect. Knowledge of the kinds of ion-pair geometries that result in large numbers of highly correlated water positions and the correct hydration of the protein is required to provide a structural basis for studying the charges in protein–water, water–water and protein–protein interactions. This information has not yet been made available even for the 1.15 Å resolution γ-crystallin structure (Basak *et al.*, 2003) because the major components of these interactions, the H atoms, are invisible in the structure. In general, H atoms are not seen using X-ray crystallography unless ultrahigh-resolution data (<1 Å) can be collected and at best their location can only be inferred by the stereochemistry around the parent atom.

The situation is different, however, for neutron crystallography (Niimura *et al.*, 1997; Bon *et al.*, 1999; Habash *et al.*, 2000; Shu *et al.*, 2000; Ostermann *et al.*, 2002), where the scattering centres are the atomic nuclei and hydrogen or deuterium (^{2}H or D) have a comparable scattering power to that of other atoms in a protein (C, N, O, S). Nevertheless, neutron scattering using hydrogenated samples can be problematic for a number of reasons: (i) incoherent scattering, predominantly from hydrogen, contributes a significant background to the scattering pattern and (ii) the negative coherent scattering length of hydrogen (-0.374×10^{-12} cm) can lead to density cancellation with adjacent atoms at medium resolution (worse than ~2.5 Å). These effects can be overcome by replacing hydrogen with deuterium, which has a much smaller incoherent cross-section (2.05 barns compared with 80.27 barns for H) and a positive coherent scattering length (0.667×10^{-12} cm), strengthening the high-resolution diffraction data and providing an order of magnitude gain in signal to noise.

From a practical view, it is possible to obtain partially deuterated crystals either by growth from D_2O solutions or by soaking crystals in D_2O solutions (Borah *et al.*, 1985; Niimura *et al.*, 1997; Bon *et al.*, 1999; Habash *et al.*, 2000; Coates *et al.*, 2001; Chatake *et al.*, 2003). This treatment leads to the replacement of some H atoms bound to nitrogen and oxygen (exchangeable H atoms) with D atoms. Unfortunately, exchangeable H atoms account for only 25% of the total H atoms; H atoms bound to C atoms are not exchangeable in this way. The only method to replace the non-exchangeable H atoms is to produce a fully deuterated protein *in vivo*. Production of significant amounts of perdeuterated protein is possible as a consequence of robust *Escherichia coli* expression systems that can be adapted to growth in D_2O and that allow high levels of protein expression in deuterated media (Vanatalu *et al.*, 1993; Gamble *et al.*, 1994; Shu *et al.*, 2000; Brockwell *et al.*, 2001; Tuominen *et al.*, 2004).

Here, we describe the production of deuterated rat γE-crystallin and the X-ray structure determination of the hydrogenated (Hγ E) and perdeuterated (Dγ E) protein in both H_2O and in D_2O at resolutions of 1.7 Å or better. To simplify discussion, we define Hγ E-H_2O as representing the hydrogenated model in H_2O solvent and Hγ E-D_2O as representing the hydrogenated model in D_2O solvent; similar constructs are defined for the deuterated proteins. The structure of native rat γE-crystallin has already been published at 2.3 Å resolution (Norledge *et al.*, 1997). Our results indicate that neither *in vivo* deuterium labelling nor crystal growth in D_2O solutions have any significant effect on the crystal structure. A new important structural detail appears in the γE-crystallin structure in a loop localized between strands *c* and *d* of the Greek-key motif 3, which is considered to be the principal structural difference between γ crystallins of low and high T_c (Norledge *et al.*, 1997).

2. Materials and methods

2.1. Production of hydrogenated and perdeuterated γE-crystallin

The plasmid pET28aγE was constructed by inserting cDNA encoding the wild-type γE-crystallin from *Rattus norvegicus* between the *Nco*I and *Not*I sites of pET28a (Novagen Inc.) For overproduction of the recombinant crystallins, BL21 (DE3) cells transformed with pET28aγE were grown in high cell-density cultures (HCDCs) using either fully hydrogenated or deuterated minimal media: 6.86 g l^{-1} $(NH_4)_2SO_4$, 1.56 g l^{-1} KH_2PO_4, 6.48 g l^{-1} $Na_2HPO_4{\cdot}2H_2O$, 0.49 g l^{-1} diammonium hydrogen citrate, 0.25 g l^{-1} $MgSO_4{\cdot}7H_2O$, 1.0 ml l^{-1} (0.5 g l^{-1} $CaCl_2{\cdot}2H_2O$, 16.7 g l^{-1} $FeCl_3{\cdot}6H_2O$, 0.18 g l^{-1} $ZnSO_4{\cdot}7H_2O$, 0.16 g l^{-1} $CuSO_4{\cdot}5H_2O$, 0.15 g l^{-1} $MnSO_4{\cdot}4H_2O$, 0.18 g l^{-1} $CoCl_2{\cdot}6H_2O$, 20.1 g l^{-1} EDTA), 5 g l^{-1} glycerol, 40 mg l^{-1} kanamycin (Meilleur *et al.*, 2004). For preparation of fully deuterated medium, hydrated mineral salts were dried out and labile H atoms were exchanged for deuterium by dissolving in a minimal volume of D_2O and then freeze-drying. Perdeuterated d_8-glycerol (CIL) was used as a carbon source. Firstly, the preparation of a deuterium-tolerant bacterial strain is necessary and was performed by a multi-stage adaptation process. Freshly transformed cells were transferred to a solid hydrogenated minimal medium. After overnight incubation at

Table 1
Crystal data, data-collection and refinement statistics for the different isotopic forms after final refinement with *SHELXL* (Sheldrick & Schneider, 1997).

Values in parentheses are for the highest resolution shell.

	HγE in H_2O solvent	HγE in D_2O solvent	DγE in H_2O solvent	DγE in D_2O solvent
Resolution (Å)	51.3–1.45 (1.54–1.45)	50.64–1.44 (1.53–1.44)	50.64–1.36 (1.45–1.36)	50.63–1.72 (1.83–1.72)
Unit-cell parameters (Å)	a = 37.2, b = 43.5, c = 102.4	a = 37.6, b = 43.4, c = 101.9	a = 37.5, b = 43.3, c = 101.6	a = 37.4, b = 43.3, c = 101.5
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Reflections	29137	29159	36431	18135
Completeness (%)	95.6 (89.8)	97.1 (90.3)	99.9 (93.7)	99.2 (89.6)
Multiplicity	3.5 (2.1)	3.7 (2.5)	7.4 (6.1)	6.9 (5.8)
$I/\sigma(I)$	7.0 (3.5)	6.1 (2.7)	5.6 (2.1)	4.7 (2.3)
R_{sym} (%)	7.6 (20.8)	8.8 (25.9)	6.7 (36.6)	8.1 (29.6)
R_{work} (%)	21.0	17.2	18.7	19.1
R_{free} (%)	25.4	21.8	21.3	22.3
DPI† (Å)	0.074	0.070	0.059	0.121
No. of atoms	1738	1768	1767	1732
No. of water molecules	121	161	145	120
No. of acetate molecules	2	2	3	2
B factors (Å^2) from Wilson plots	14.3	12.5	11.2	16.6
All atoms	21.1	18.1	16.9	22.2
Protein	20.3	16.9	15.9	21.5
Waters	29.7	28.2	26.3	30.6
Acetate	43.0	55.4	31.3	39.3
R.m.s.d bonds (Å)	0.016	0.017	0.017	0.018
R.m.s.d angles (°)	1.689	1.766	1.768	1.761
Ramachadran plot (%): most favoured/additional allowed	90.6/9.4	89.9/10.1	91.3/8.7	89.9/10.1

† DPI is the diffraction-component precision index (Cruickshank, 1999).

303 K, cells were plated on fully deuterated minimal medium. Following incubation for 30 h at 303 K, cells were transferred to a flask containing 10 ml fully deuterated minimal medium; this overnight culture was again incubated at 303 K. The adaptation to growth in D_2O was established after four cycles of dilution and growth in 10 ml deuterated minimal medium. 1.5 l of deuterated medium was inoculated with 100 ml preculture in a 3 l fermenter (Labfors, Infors). During the batch and fed-batch phases the pH was adjusted to 6.9 (by the addition of NH_3 or NaOD as appropriate) and the temperature was adjusted to 303 K to allow a more efficient control of the growth rate. The gas-flow rate of sterile filtered air was $\geq$0.5 l min^{-1}. Stirring was adjusted to ensure a dissolved oxygen tension (DOT) of 30%. The fed-batch phase was initiated when the optical density at 600 nm reached 6.0. Hydrogenated or deuterated glycerol was added to the culture to keep the growth rate stable during fermentation. When OD_{600} reached ~12, γE-crystallin overexpression was induced by the addition of 1 m*M* IPTG and incubation continued for 24 h. Cells were then harvested, washed with 10 m*M* HEPES pH 6.4 and stored at 193 K.

2.2. Protein purification

Hydrogenated and perdeuterated γE-crystallins were purified in hydrogenated buffers using similar protocols. Briefly, 10 g of frozen cell paste was resuspended in 150 ml lysis buffer, 100 m*M* HEPES pH 6.4, 5 m*M* DTT and EDTA-free anti-proteases (Roche). Cells were lysed using the cellD system (Constant Systems) at 130 MPa and centrifuged at 288 K for 45 min (40 000*g*). The supernatant was loaded onto a 10 ml SP Sepharose column (Amersham Biosciences) equilibrated with lysis buffer. The protein was eluted with a gradient of NaCl (0–0.3 *M*). Fractions containing crystallins were pooled and precipitated with 70% $(NH_4)_2SO_4$. After centrifugation, the pellet was resuspended in 3 ml lysis buffer and loaded onto a Superdex 200 Hi-load prep-grade gel-filtration column (Amersham Biosciences) equilibrated with lysis buffer. Crystallin-containing fractions were pooled, concentrated and dialysed against water for 16 h. After treatment with 5 m*M* DTT for 30 min, the pure crystallin was concentrated again. Typical yields were 4–5 mg of pure protein per gram of hydrogenated or deuterated cell paste.

The level of isotopic incorporation was determined by electrospray ionization mass spectrometry. Hydrogenated and deuterated proteins were loaded onto a Protein MacroTrap cartridge (Michrom) and desalted with H_2O supplemented with 0.1% TFA. Samples were eluted with 70% CH_3CN/H_2O (9:1 by volume), 0.1% TFA and analyzed by ESI–MS. Mass determinations were performed on an API III+ triple quadrapole mass spectrometer equipped with an ion-spray source (PE Sciex). The electrospray probe tip was held at 5 kV and the declustering potential was set at 80 V. Data were processed using *MASSPEC* v.3.3 software and the reconstructed molecular-mass profiles were obtained from the m/z mass spectrum using a deconvolution algorithm. The calibration of the spectrometer was carried out using a polypropylene glycol solution. ESI–MS were recorded for hydrogenated and deuterated γE proteins, resulting in a molecular weight of 21 132 and 22 113 kDa, respectively, in hydrogenated buffers. γE-crystallin has a theoretical molecular weight of 21 262 kDa for its hydrogenated form. The experimental weight was

131 Da lower, indicating the loss of the N-terminal methionine by the bacterial aminopeptidase. This is a common post-translational modification for both prokaryotic and eukaryotic proteins, especially when the second position in the polypeptide chain is occupied by a small side-chained amino acid such as glycine, which is the case for most γ-crystallins (Hirel *et al.*, 1989). The theoretical weight of D-γE is 21.262 + [1002 (number of non-exchangeable deuteriums) × 1.006 (the mass difference between deuterium and hydrogen)] − 139 (deuterated N-terminal methionine) = 22 131 kDa. The level of deuteration was therefore estimated to be 98.2%, taking into account the loss of the deuterated N-terminal methionine. To prepare H-γE and D-γE-crystallin in deuterated solvent, proteins were diluted and concentrated five times against D_2O pD 7.4, which corresponds to an original protonic pH of 7.0 (Lumry *et al.*, 1951).

2.3. Crystallization and X-ray diffraction data collection

H- and D-γE-crystallins were crystallized from both hydrogenated and deuterated crystallization solutions using an adaptation of the native protein crystallization protocol (Norledge *et al.*, 1997). Crystals for X-ray diffraction studies were grown using the hanging-drop vapour-diffusion method. For each batch of protein, conditions were optimized at 293 K by varying the amount of polyethylene glycol (PEG) 3350 precipitant with 0.1 *M* magnesium acetate pH 7.4 (pD 7.8) in the reservoir solution. Deuterated magnesium acetate was prepared by recrystallizing hydrogenated salts in D_2O. The PEG stock solution was made by dissolving PEG of the appropriate molecular weight in D_2O. Sharp-edged crystals grew after a few days in 4 µl drops formed of equal volumes of the protein solution (10 mg ml^{-1}) and 18%(*w*/*v*) (20% for deuterated crystallization condition) PEG 3350 dissolved in 0.1 *M* magnesium acetate pH 7.4. For cryoprotection, the crystals were soaked in five steps in buffers consisting of mother liquor with an increasing concentration of H_8-glycerol or D_8-glycerol to 0–20% and then flash-frozen in a stream of gaseous nitrogen at 100 K. X-ray diffraction data were collected on beamline ID14-EH4 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. The diffraction images were indexed and integrated with the programs *MOSFLM* (Leslie *et al.*, 1986) or *XDS* (Kabsch, 1993) and the resulting intensities were scaled and merged using the program *SCALA* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Data-collection statistics and final structural refinement statistics are summarized in Table 1.

2.4. Structure determination, refinement, comparison and analysis

The 2.3 Å γE-crystallin structure (chain *A*, without water molecules) determined at 300 K (PDB code 1a5d; Norledge *et al.*, 1997) was used as a starting model to independently refine the four isotopic forms of the γE-crystallin. In each isotopic crystal form the crystallographic asymmetric unit contains one crystallin molecule, compared with the two molecules in 1a5d (space group $P2_1$). The three last C-terminal amino acids (Asp171-Phe172-Tyr173) were removed from the models owing to their high flexibility. The refinements were performed successively using *REFMAC* without TLS refinement (v.5; Murshudov *et al.*, 1997) from the *CCP*4 package and *SHELXL* with the *SHELXPRO* interface (Sheldrick & Schneider, 1997) and the initial solvent model was created using *ARP*/*wARP* v.5.0 (Lamzin & Wilson, 1993). Water molecules were added to the coordinate list according to the following criteria: (i) a water molecule must make at least one stereochemically relevant hydrogen bond, (ii) it should have well defined density above 1σ in a $2F_o - F_c$ map, (iii) it should have density above 3σ in an $F_o - F_c$ map and (iv) its *B* factor should be less than 60 Å^2. The starting model was subjected to rigid-body refinement with only the *A* chain; it was then subjected to alternate runs of positional refinement, simulated-annealing and individual *B*-factor refinement with the addition of water molecules and with intermittent manual model corrections using the program *O* (Jones *et al.*, 1991). In particular, manual adjustments of the loop between the *c* and *d* strands in motif 3 (residues 115–122) were made as this loop differs significantly from the starting model. In the models, all of which are at a significantly higher resolution than the previously reported structure (Norledge *et al.*, 1997), the *cd* loop showed a clear double conformation for both the main chain and the side chains. The final models contained water molecules, all of which refined at unit occupancy except for molecules close to the *cd* loop, and two (three for DγE-H_2O) acetate molecules. The models were also subjected to quality analysis during the various refinement stages by the use of omit maps. Refinement statistics are presented in Table 1. The root-mean-square deviations (r.m.s.d.s) were calculated using the program *LSQKAB* from *CCP*4 using the main chain and the side chains of the different models (and chain *A* for the 1a5d model). Secondary structures were assigned using *PROCHECK* (Laskowski *et al.*, 1993).

3. Results

3.1. Crystal structures of H- and D-γE-crystallins in H_2O and D_2O solutions

The four different γE-crystallin structures were determined at resolutions between 1.72 and 1.36 Å. Table 1 provides the crystallographic parameters observed for each structure. The $2F_o - F_c$ electron-density maps at 1.0σ for each protein were well defined for all the polypeptide main-chain atoms, with the exception of a loop localized between residues 116–119 and the C-terminus (residues 171–173); this localized disorder was also observed in native rat γE-crystallin.

The X-ray analysis of all the forms of rat γE-crystallin has shown that the molecules are organized into two similar globular domains related by a pseudo-dyad. Each domain is constructed from about 80 amino acids arranged as two 'Greek-key' motifs (Blundell *et al.*, 1981; Bax *et al.*, 1990), which are again related by a pseudo-dyad, leading to a symmetrical molecule. Two Greek-key motifs, each containing

Table 2
Comparison of amino-acid sequences comprising the *cd* loop for high-T_c and low-T_c γ-crystallins (adapted from Norledge *et al.*, 1997).

Group	Protein name and organism	Sequence
High transition temperature (high T_c)	Bovine γE	115--F--I---F--128
	Bovine γF	115--F--I---F--128
	Rat γE	115--F--F---F--128
	Rat γF	115--F--F---F--128
Low transition temperature (low T_c)	Bovine γB	115--L--V---L--128
	Bovine γD	115--F--I---L--128
	Human γC	115--L--I---L--128
	Rat γA	115--L--I---M--128
	Rat γB	115--L--I---L--128
	Rat γC	115--L--V---L--128

four antiparallel β-strands (*a*, *b*, *c*, *d*), contribute to two antiparallel β-pleated sheets (*badc*′) in which three strands, *b*, *a* and *d*, of one motif are hydrogen-bonded to strand *c*′ of the second motif. The residues with side chains between the β-sheets and the motifs are hydrophobic. The connection between the N- and the C-terminal domains is achieved by a random-coil sequence encompassing residues 81–87. This overall topology, which has been observed in previous studies, is confirmed by our results here.

3.2. The *cd* loop in the Greek-key motif 3

The most striking difference between the models presented here and the previously determined 2.3 Å structure (Norledge *et al.*, 1997) is found in the *cd* loop (residues 115–122) of the Greek-key motif 3 (Table 2). In the native structure the electron-density map for this loop was not well defined for one conformation. In the present model, however, non-continuous electron density is observed, indicating a certain flexibility. It has been possible to build two conformations of the loop in the non-continuous electron-density map in each isotopic form owing to the low occupancy of the two conformations (see Fig. 1).

Interestingly, in each structure the crystal packing is maintained by a loop formed by Phe116, His117, Phe118, Ser119, Asp120, Phe121 and His124, previously characterized as flexible and

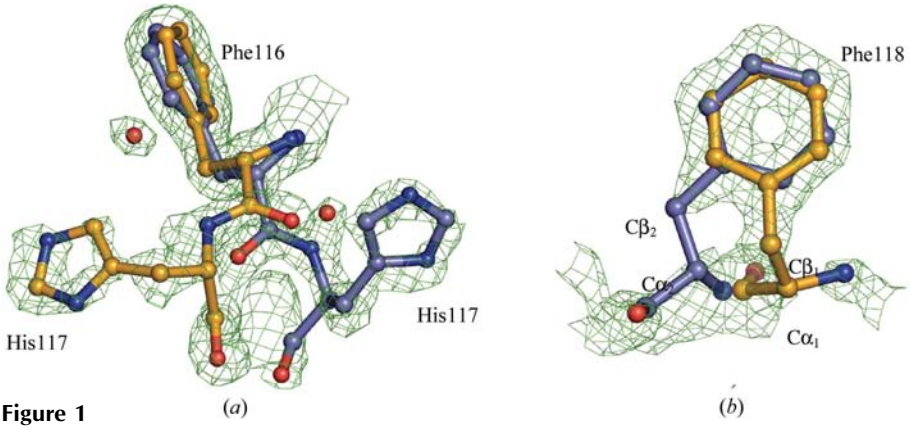

Figure 1
$2F_o - F_c$ electron-density map of the *cd* loop of Greek-key motif 3 of DγE-H_2O. (*a*) Representation of the double conformation of residues Phe115 and His166 with the electron density at 1σ. (*b*) The aromatic rings of the Phe117 residues are located in the same position but the C^{α} and C^{β} atoms are displaced. All C atoms of the loop are shown in ball-and-stick representation in light yellow for the first conformation and in purple in the second conformation. N atoms are in blue and O atoms are in red. The figures were prepared using *PyMOL* (DeLano, 2002).

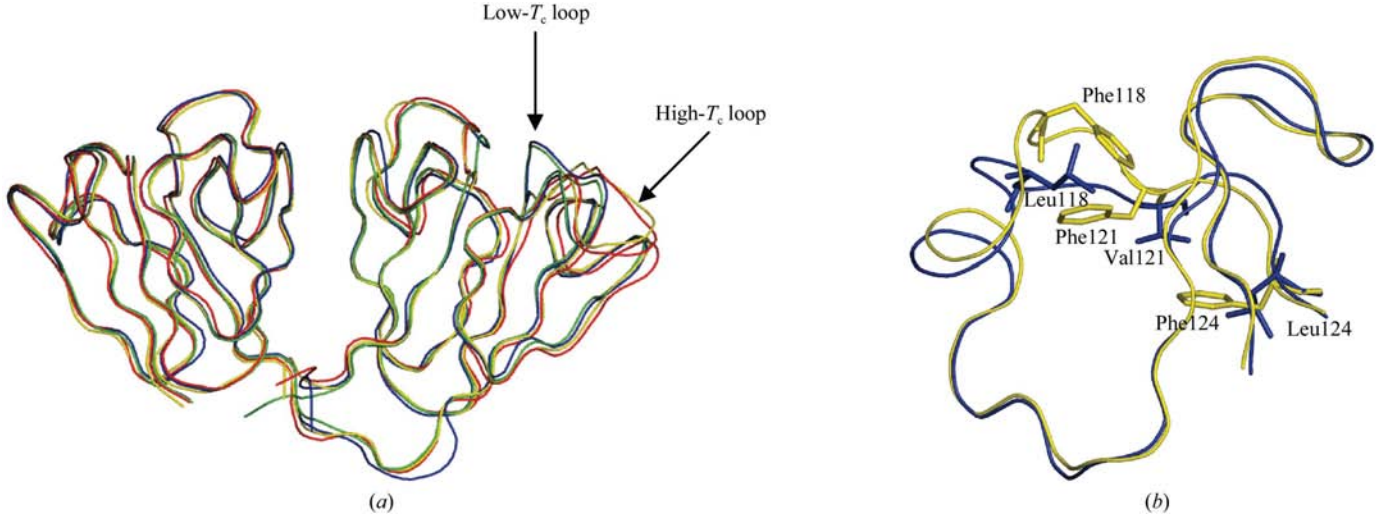

Figure 2
Comparison of γ-crystallin structures. (*a*) Superposition of the structure of the high-T_c proteins γF-crystallin from rat (1a45; red) and the 2.3 Å γE-crystallin from rat (1a5d; yellow) and the low-T_c proteins bovine γB-crystallin (1amm; blue) and bovine γD-crystallin (1elp; green). The r.m.s. deviations on C^{α} atoms between a low-T_c crystallin (bovine γB) and the superimposed structures are 0.43 Å for the other low-T_c crystallin (bovine γD) and 0.69 and 0.88 Å for the 1a5d and 1a45 high-T_c crystallins. The structures are shown as tube representations using the program *PyMOL* (DeLano, 2002). (*b*) Close-up of the loop between strands *c* and *d* in motif 3. The C^{α} backbone of the motif 3 from the bovine low-T_c γB (in blue) is superposed on the motif 3 from the rat high-T_c γE-crystallin (in yellow). The hydrophobic core is bulkier in the high-T_c protein, where aromatic residues such as Phe118, Phe121 and Phe124 push the loop outside of the molecule.

Table 3
Main-chain, side-chain and all-atoms r.m.s. deviations (in Å), respectively, between all rat γE-crystallin structures.

The hydrogenated (H) and deuterated (D) γE-crystallin models in hydrogenated solvent (H_2O) and in deuterated solvent (D_2O) and the chain *A* of the native model (1a5d) are compared without the second conformation of the *cd* loop (116–120) and the C-terminal end (172–174) taken into account.

Model	HγE-D_2O			DγE-H_2O			DγE-D_2O			1a5d		
HγE-H_2O	0.16	0.44	0.31	0.15	0.42	0.30	0.14	0.40	0.28	0.30	0.89	0.61
HγE-D_2O				0.07	0.35	0.23	0.10	0.40	0.27	0.34	0.91	0.64
DγE-H_2O							0.08	0.37	0.24	0.33	0.94	0.65
DγE-D_2O										0.33	0.93	0.65

variable in the γ-crystallin family (Norledge *et al.*, 1997). In the present models, this loop interacts with a symmetry-equivalent molecule *via* hydrogen-bonding and ionic interactions. The part of the protein presents two different conformations in which main chain and side chains are separated by a few angstroms, in particular His117, which is separated by 4.47 Å for C^{α} and 11.10 Å for C^{ε} (see Fig. 2*a*). His117 in the first conformation makes primarily intramolecular interactions with Gln100 and some water molecules, whereas in the second conformation His116 interacts mainly with residues of a symmetry equivalent molecule and causes a shift in the position of several subsequent amino acids. The two conformations of Phe118 have the same positioning in space of the aromatic cycle in spite of a divergence in the localization of C atoms C^{α} and C^{β} (separated by 2.12 and 2.72 Å, respectively; Fig. 2*b*). The shift of the main chain is compensated for by double conformations of Ser119 and Asp120, which mean that Phe121 is positioned in the same way in both conformers. In the second conformation, His122 is in a different conformation owing to the proximity of Asp12. The aromatic ring of His122 is not perpendicular to the plane of the side chain of Asp120: they are 'parallel' in order to limit the steric effects generated by this new conformation. Owing to this second conformation, a large number of hydrogen bonds are established between His122 and His118 of a symmetry-equivalent molecule and so crystal packing is favoured.

3.3. Comparison between the different isotopic forms of γE-crystallin

The X-ray structures of H-γE- and D-γE-crystallin in hydrogenated and deuterated water between 1.36 and 1.72 Å are very similar to the native γE-crystallin at 2.3 Å resolution (Norledge *et al.*, 1997): each model presents the same fold. The r.m.s. difference between the models (ignoring the second conformation of the *cd* loop and the C-terminal three residues) are given in Table 3. The main-chain and side-chain atoms do not deviate very much from one model to another, with an average r.m.s. difference that is never higher than 0.435 Å (Table 3). The deviations with respect to the 1a5d (chain *A*) structure are somewhat higher than those between the four structures reported here.

The reason for the drop in the r.m.s. difference values is most probably the lower resolution of the native model at 2.3 Å. Graphical representations of the main-chain and side-chain r.m.s. deviations for each isotopic form are presented in Fig. 3, where HγE-H_2O was taken (arbitrarily) as a reference. The majority of regions present the same well ordered fold; consequently, the deviation is not significant for the bulk of the residues. The highest r.m.s. deviations correspond to the C-terminal part (residues 171–173) and the *cd* loop (residues 115–121), which reach 4.1 Å. Electron density in the C-terminal region is not well defined and this part was previously characterized as being highly flexible (Najmudin *et al.*, 1993). For the *cd*-loop region, electron density is poorly defined for each isotopic form, although it suggests the presence of a double conformation. The native structures present some deviations, especially for the side chains, but this model was obtained from crystals grown under different crystallization conditions, with a different space group and also with different crystal contacts (see Fig. 3*b*). All of these factors should be taken into account to explain these small deviations. Except for these minor deviations, all the structures, whether native or recombinant, hydrogenated or deuterated or in a hydrogenated or deuterated solvent, can be considered as being highly similar.

3.4. Comparison of *B* factors

Analysis of the *B* values as a function of residue number shows that the overall structure of the proteins is relatively rigid (see Fig. 4 and Table 4). However, disorder in the region of the *cd* loop, with its double conformation, is less visible from the *B*-factor plot than for the C-terminal part of the protein. The *B* factors of the polypeptide chain are particularly high in this last region and are quite different for each of the isotopic forms. *B* factors derived from the Wilson plots are slightly lower than those obtained after the final refinements (see Table 1), but show the same trends. Values for the two homogeneous systems (HγE-H_2O and DγE-D_2O) are higher than those of the two heterogeneous systems (HγE-D_2O and DγE-H_2O). Comparing the two systems in which the hydrogenated protein is present, an effect of solvent is observable. Indeed, the system in D_2O presents an average *B* factor that is significantly smaller than that in H_2O for main-chain, side-chain and residue *B* factors (see Table 4). On the other hand, if the *B* factors of hydrogenated and deuterated proteins in H_2O are compared, an effect on the flexibility of the protein generated by the perdeuteration is also visible. Indeed, the average *B* factor of hydrogenated protein in H_2O is higher than that of deuterated protein in H_2O. However, when the two deuterated proteins are compared, the *B* factors of deuterated protein in D_2O are higher than those of deuterated protein in H_2O. This latter surprising result may, however, arise from either the different resolution of the data (1.36 Å for DγE-H_2O and 1.72 Å for DγE-D_2O) or the poorer quality of the data for deuterated protein in D_2O (see Table 1). In order to check this we also carried out refinement of all structures at 1.72 Å resolution (*i.e.* that of the lowest resolution data set). Although the DγE-D_2O structure still had the largest thermal parameters, those of the other structures were

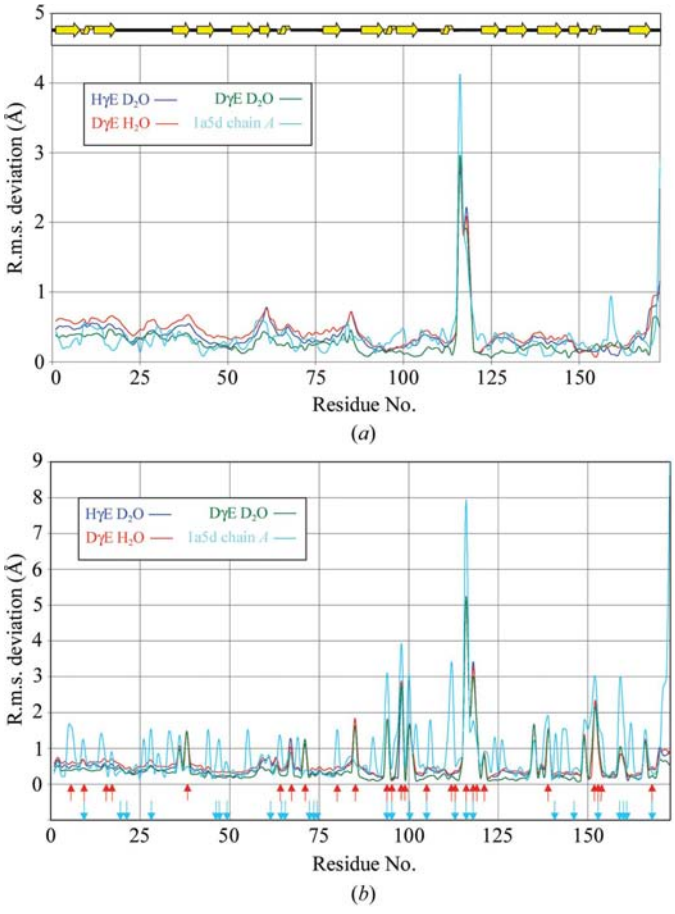

Figure 3
R.m.s. deviations of (*a*) the main chains and (*b*) the side chains between HγE-H_2O, taken arbitrarily as a reference, and the other forms. Above the first plot there is a schematic representation of the secondary structure predicted by the program *PROCHECK* (Laskowski *et al.*, 1993) and below the second plot there is a schematic representation of the crystalline contacts. Each amino acid engaged in a crystal contact is represented by an arrow: up arrows indicate the HγE-H_2O model and down arrows indicate the 1a5d model.

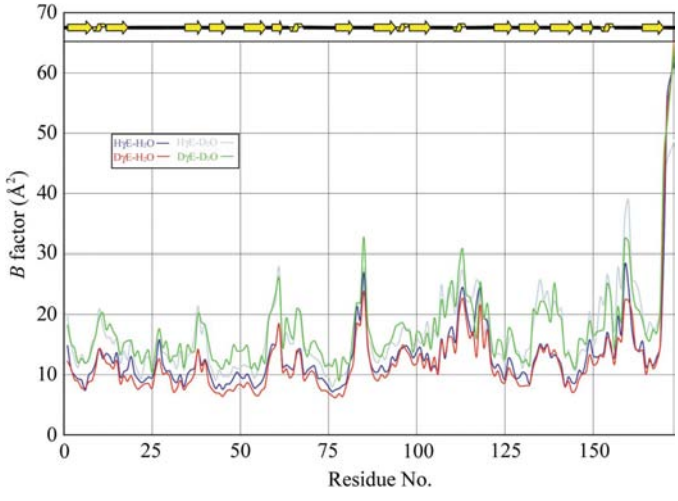

Figure 4
B factors of main-chain atoms plotted against residue number of the different models after *REFMAC* refinement (Murshudov *et al.*, 1997).

Table 4
B factors (Å^2) of main chain, side chains and residues of each form of γE-crystallin after *REFMAC* refinement (Murshudov *et al.*, 1997) from *CCP*4.

	HγE-H_2O	HγE-D_2O	DγE-H_2O	DγE-D_2O
Main chain	17.0	13.4	12.4	17.6
Side chains	20.5	17.3	16.5	22.3
Residues	19.8	16.2	15.2	21.0

not significantly lower. This was also confirmed by carrying out refinement with both *SHELXL* and *REFMAC* (data not shown). Quantitative comparisons concerning DγE-D_2O should therefore be treated with caution.

4. Discussion

The refined crystal structure of hydrogenated and perdeuterated forms of rat γE-crystallin in H_2O and D_2O have been presented at resolutions up to 1.36 Å and have been compared with the previously determined structure of the γE-crystallin protein at 2.3 Å. The structures are remarkably similar, sharing a high degree of internal symmetry which is common to all γ-crystallins. At this resolution, the major observed difference with the previous native model is the existence of a double conformation of the loop between strands *c* and *d* in Greek-key motif 3. One of these conformations is identical to that found previously and the other differs in the path taken by the main chain whilst maintaining the position of the phenylalanine side chains. Norledge *et al.* (1997) and White *et al.* (1989) concluded that different amino-acid compositions allow different conformations of the *cd* loop that are characteristic of the phase-transition group (see Table 2). In the high-T_c group the loop is composed of large residues such as phenylalanine in positions 118, 121 and 125, whereas in the low-T_c group smaller amino acids are observed at these

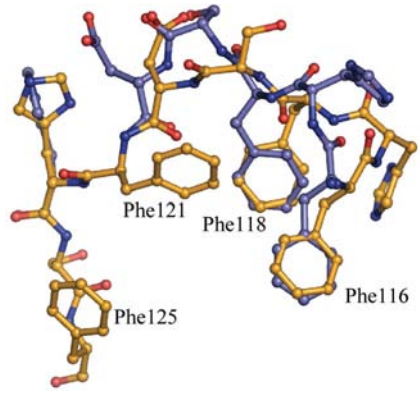

Figure 5
Ball-and stick representation of the hydrophobic core maintained by phenylalanines 116, 118, 121 and 125 in the two conformations of the *cd* loop. Despite the double conformation, aromatic residues present their side chains at approximately the same position. All C atoms of the loop are shown in light yellow for the first conformation and in purple for the second conformation. N atoms are in blue and O atoms in red. The figures were prepared using *PyMOL* (DeLano, 2002).

positions. The hydrophobic core formed between the Greek-key motifs 3 and 4 is less tightly packed in this last group. Although the present models of rat γE-crystallin show two well defined conformations of the *cd* loop, which is the major structural difference between high- and low-T_c γ-crystallins, both conformations form a hydrophobic core consistent with high-T_c structure. In the second conformation, His117 plays a major role in the crystal packing by interacting with a symmetry equivalent and the phenylalanines 116, 118, 121 and 125 conserve the same position (see Figs. 2*b* and 5). The bulk hydrophobic core formed mainly by these latter residues is still preserved. Owing to this conservation of the positions of the hydrophobic residues, we do not consider these two conformations to be representative of the two phase-separation groups. In addition, the deviation between the second conformation of the loop and the low-T_c loop of bovine γB crystallin for example (Kumaraswamy *et al.*, 1996) reaches 5.8 Å (data not shown). Thus, the flexibility of this loop is necessary to maintain both the integrity of the hydrophobic core and to allow interactions with a symmetry-equivalent molecule through histidine residues. This small flexibility of the *cd* loop favours *in vivo* the specific interaction between the γE/F crystallins and the major intrinsic protein (MIP; Fan *et al.*, 2005). It could play a major role in the formation of the water channel and in the ability of MIP subunits to form intercellular junctions maintaining the organization and the cell physiology (Ball *et al.*, 2003; Gonen *et al.*, 2004).

Perdeuteration of proteins is a powerful method for improving protein structure determination by NMR (Katz & Crespi, 1966) and neutron crystallography (Gutberlet *et al.*, 2001). One of the major advantages of a neutron crystallographic study of γ-crystallins would be to make a detailed description of the solvent organization and hydrogen-bonding network around the molecule, which would provide a structural basis for studying protein–water, water–water and protein–protein interactions. An obvious requirement for the validity of isotopic substitution in neutron crystallography is that the hydrogenated and the deuterated protein do not differ significantly in their crystal structure.

In this study, the different isotopic forms of the rat γE-crystallin have been shown to possess no major structural differences (r.m.s. deviation of 0.27 Å between all isotopic forms) as has been observed in previous studies on partial or fully deuterated systems (Tuominen *et al.*, 2004; Wlodawer *et al.*, 1986, 1989; Coates *et al.*, 2001). Thus, perdeuteration does not effect the position of atoms in space, but there are some effects from the isotopic change of non-exchangeable and exchangeable hydrogen with D atoms. Interestingly, recent comparative FTIR studies of hydrogenated and perdeuterated P450cam-(C334A) show that the structure and dynamics of the proteins under ambient conditions of temperature and pressure are very similar (Meilleur *et al.*, 2004).

It has been demonstrated that bonds involving deuterium (instead of hydrogen) are stronger; thus, the frequency of vibration of these bonds is lower (Benjamin & Benson, 1962) and thus thermal *B* factors should be reduced. A deuterium bond will be less flexible than a hydrogen bond, thus D_2O and perdeuteration increase the rigidity of the native structure of proteins (Cioni & Strambini, 2002). It should be noted that the presence of D_2O as solvent rather than perdeuteration should have the most striking effect as it is hydrogen (deuterium) bonds which are primarily responsible for the stability of secondary-structure elements of a protein. Our studies on γE-crystallin have shown that differences in *B* factors are minimal and do not allow us to make such quantitative comparisons. Lutter *et al.* (1993) have used the characteristic of increased rigidity in D_2O to aid the purification and subsequent crystallization of the F_1-ATPase from bovine heart mitochondria. However, previous studies have demonstrated that the stability and the enzymatic activity of a protein can be affected quite significantly by perdeuteration for unstable proteins (Brockwell *et al.*, 2001). This is not the case for the highly stable γE-crystallin protein, which has similar transition temperatures for hydrogenated and deuterated forms in H_2O and in D_2O (in preparation).

In summary, our data, together with previous results from the literature, indicate that perdeuterated proteins have a similar structure to their hydrogenated counterparts. Analysis of thermal *B* factors suggests that in the case of γE-crystallin, which is already an exceptionally stable protein, these parameters do not indicate a significant stabilizing effect of deuterium. Structural information obtained using neutron diffraction will therefore be representative of the hydrogenated protein packing in γ-crystallin crystals and can be used as a structural basis for analysing protein–protein interactions

We are grateful to N. Lubsen (Nijmegen) for providing the initial γE-crystallin clone and C. Slingsby (Birkbeck College, London) for illuminating discussions on crystallin structure. We also thank Marie-Thérese Dauvergne (EMBL Grenoble) for help with cloning, expression and purification of the protein, the staff of the ILL/EMBL Deuteration Laboratory, Grenoble for assistance, David Lemaire (IBS, France) for the mass-spectrometry analysis, Ingar Leiros and Gordon Leonard (ESRF, Grenoble) for help in data collection and Susana Teixeira (ILL) for data treatment. This project has benefitted from the activities of the EU-funded DLAB project under contracts HPRI-CT-2001-50035 and RII3-CT-2003-505925 as well as activities funded under EPSRC grant GR/R99393/01.

References

Ball, L. E., Little, M., Nowak, M. W., Garland, D. L., Crouch, R. K. & Schey, K. L. (2003). *Invest. Ophthalmol. Vis. Sci.* **44**, 4820–4828.
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.
Benedek, G. (1971). *Appl. Opt.* **10**, 459–473.
Benedek, G., Clark, J., Serrallach, E., Young, C., Mengel, L., Sauke, T., Bagg, A. & Benedek, K. (1979). *Philos. Trans. R. Soc. London A*, **293**, 329–340.

Benjamin, L. & Benson, G. (1962). *J. Phys. Chem.* **67**, 858–861.
Berland, C. R., Thurston, G. M., Kondo, M., Broide, M. L., Pande, J., Ogun, O. & Benedek, G. B. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 1214–1218.
Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. & Wistow, G. (1981). *Nature (London)*, **289**, 771–777.
Bon, C., Lehmann, M. S. & Wilkinson, C. (1999). *Acta Cryst.* D**55**, 978–987.
Borah, B., Chen, C. W., Egan, W., Miller, M., Wlodawer, A. & Cohen, J. S. (1985). *Biochemistry*, **24**, 2058–2067.
Brockwell, D., Yu, L., Cooper, S., McCleland, S., Cooper, A., Attwood, D., Gaskell, S. J. & Barber, J. (2001). *Protein Sci.* **10**, 572–580.
Broide, M. L., Berland, C. R., Pande, J., Ogun, O. O. & Benedek, G. B. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 5660–5664.
Chatake, T., Ostermann, A., Kurihara, K., Parak, F. G. & Niimura, N. (2003). *Proteins*, **50**, 516–523.
Chirgadze, Y., Driessen, H., Wright, G., Slingsby, C., Hay, R. E. & Lindley, P. F. (1996). *Acta Cryst.* D**52**, 712–721.
Cioni, P. & Strambini, G. B. (2002). *Biophys. J.* **82**, 3246–3253.
Clark, J. I. & Carper, D. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 122–125.
Coates, L., Erskine, P. T., Wood, S. P., Myles, D. A. & Cooper, J. B. (2001). *Biochemistry*, **40**, 13149–13157.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.
DeLano, W. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA.
Delaye, M. & Tardieu, A. (1983). *Nature (London)*, **302**, 415–417.
Fan, J., Fariss, R. N., Purkiss, A. G., Slingsby., C., Sandilands, A., Quinlan, R., Wistow, G. & Chepelinsky, A. B. (2005). *Mol. Vis.* **11**, 76–87.
Gamble, T. R., Clauser, K. R. & Kossiakoff, A. A. (1994). *Biophys. Chem.* **53**, 15–25.
Gonen, T., Sliz, P., Kistler, J., Cheng, Y. & Walz, T. (2004). *Nature (London)*, **429**, 193–197.
Gutberlet, T., Heinemann, U. & Steiner, M. (2001). *Acta Cryst.* D**57**, 349–354.
Habash, J., Raftery, J., Nuttall, R., Price, H. J., Wilkinson, C., Kalb, A. J. & Helliwell, J. R. (2000). *Acta Cryst.* D**56**, 541–550.
Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G. & Blanquet, S. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 8247–8251.
Ishimoto, C., Goalwin, P., Sun, S., Nishio, I. & Tanaka, T. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 4414–4416.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Katz, J. J. & Crespi, H. L. (1966). *Science*, **151**, 1187–1194.
Kumaraswamy, V. S., Lindley, P. F., Slingsby, C. & Glover, I. D. (1996). *Acta Cryst.* D**52**, 611–622.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–149.
Laskowski, R., MacArthur, M., Moss, D. & Thornton, D. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W., Brick, P. & Wonacott, A. T. (1986). *Daresbury Lab. Inf. Quart. Protein Crystallogr.* **18**, 33–39.
Lumry, R., Smith, E. L. & Glantz, R. R. (1951). *J. Am. Chem. Soc.* **73**, 4330–4340.
Lutter, R., Abrahams, J. P., van Raaij, M. J., Todd, R. J., Lundqvist, T., Buchanan, S. K., Leslie, A. G. & Walker, J. E. (1993). *J. Mol. Biol.* **229**, 787–790.
Meilleur, F., Contzen, J., Myles, D. A. & Jung, C. (2004). *Biochemistry* **43**, 8744–8753.
Murshudov, G. N., Vagin, A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Najmudin, S., Nalini, V., Driessen, H., Slingsby, C., Blundell, T., Moss, D. & Lindley, P. (1993). *Acta Cryst.* D**49**, 223–233.
Niimura, N., Minezaki, Y., Nonaka, T., Castagna, J. C., Cipriani, F., Hoghoj, P., Lehmann, M. S. & Wilkinson, C. (1997). *Nature Struct. Biol.* **4**, 909–914.
Norledge, B. V., Hay, R. E., Bateman, O. A., Slingsby, C. & Driessen, H. P. (1997). *Exp. Eye Res.* **65**, 609–630.
Ostermann, A., Tanaka, I., Engler, N., Niimura, N. & Parak, F. G. (2002). *Biophys. Chem.* **95**, 183–193.
Philipson, B. (1969). *Invest. Ophthalmol.* **8**, 258–270.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Shu, F., Ramakrishnan, V. & Schoenborn, B. P. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3872–3877.
Siezen, R. J., Fisch, M. R., Slingsby, C. & Benedek, G. B. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 1701–1705.
Siezen, R. J., Wu, E., Kaplan, E. D., Thomson, J. A. & Benedek, G. B. (1988). *J. Mol. Biol.* **199**, 475–490.
Tanaka, T., Ishimoto, C. & Chylack, L. T. Jr (1977). *Science*, **197**, 1010–1012.
Tardieu, A., Veretout, F., Krop, B. & Slingsby, C. (1992). *Eur. Biophys. J.* **21**, 1–12.
Thomson, J. A., Schurtenberger, P., Thurston, G. M. & Benedek, G. B. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 7079–7083.
Tuominen, V. U., Myles, D. A., Dauvergne, M. T., Lahti, R., Heikinheimo, P. & Goldman, A. (2004). *Acta Cryst.* D**60**, 606–609.
Vanatalu, K., Paalme, T., Vilu, R., Burkhardt, N., Junemann, R., May, R., Ruhl, M., Wadzack, J. & Nierhaus, K. H. (1993). *Eur. J. Biochem.* **216**, 315–321.
White, H. E., Driessen, H. P., Slingsby, C., Moss, D. S. & Lindley, P. F. (1989). *J. Mol. Biol.* **207**, 217–235.
Wlodawer, A., Borkakoti, N., Moss, D. & Howlin, B. (1986). *Acta Cryst.* B**42**, 379–387.
Wlodawer, A., Savage, H. & Dodson, G. (1989). *Acta Cryst.* B**45**, 99–107.