Acta Cryst. (1993). D**49**, 223–233

Structure of the Bovine Eye Lens Protein γB(γII)-Crystallin at 1.47 Å

BY S. NAJMUDIN AND V. NALINI

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

H. P. C. DRIESSEN

Imperial Cancer Research Fund Unit of Structural Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

C. SLINGSBY

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

T. L. BLUNDELL

Imperial Cancer Research Fund Unit of Structural Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

D. S. MOSS

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

AND P. F. LINDLEY

SERC Daresbury Laboratory, Warrington WA4 4AD, England

(*Received* 6 *April* 1992; *accepted* 14 *July* 1992)

Abstract

The molecular structure of calf γB-crystallin (previously called γII), a lens-specific protein, has been refined to a crystallographic *R* factor of 18.1% for all reflection data, between 8.0 and 1.47 Å, 25 959 *hkl* measured at 293 (1) K. 230 water molecules have been defined by difference Fourier techniques and included in a restrained least-squares refinement. Difference Fourier maps clearly indicated the presence of multiple sites for the sulfur atoms of Cys 18 and Cys 22 which were therefore given coupled second-site occupancies during the refinement. The sulfur atom in the major position of Cys 22 is in the reduced state. Either of the Cys 18 sites can form a high-energy disulfide bridge with the minor position of Cys 22. The position of the carboxy terminus and many other surface side chains have been further defined including the RGD signal peptide. The hydration of the backbone and the interdomain region has been analysed. 27 water molecules make extensive contacts to a single protein molecule and thus contribute to its stability.

1. Introduction

The eye lens fibre cells are packed with globular proteins, α-, β- and γ-crystallins (Wistow & Piatigorsky, 1988), in a non-periodic array (Delaye & Tardieu, 1983) such that the appropriate refractive index is reached while light scattering is minimized (Benedek, 1984). A gradient of protein concentration along the optical axis is formed in some species which is thought to minimize spherical aberration (Fernald & Wright, 1983). The various crystallin families and their individual members are differentially expressed during development (Lubsen, Aarts & Schoenmakers, 1988; Piatigorsky, 1989) leading to different mixtures of crystallins along the optical axis and along the protein gradient. γ-Crystallins are a family of closely related proteins of molecular weight 20 000 dalton that are enriched in those regions of eye lenses with the highest refractive index (Slingsby, 1985) and are thought to have a structural role in facilitating the relatively low water levels (Blundell *et al.*, 1981).

Low-resolution X-ray diffraction studies have revealed the two-domain structure of γ-crystallin (Blundell *et al.*, 1978). These proteins have high internal symmetry as each domain is comprised of two 'Greek key' motifs related by a pseudodyad and the N- and C-terminal domains are related by a further approximate dyad (Blundell *et al.*, 1981; Chirgadze, Sergeev, Fomenkova & Oreshin, 1981). The importance of the linker region to the organi-

0907-4449/93/020223-11$06.00

zation of the domains and the formation of symmetrical oligomers was revealed when the structure of a related crystallin, βB2 homodimer, was solved by molecular replacement (Bax *et al.*, 1990; Lapatto *et al.*, 1991). Although the interface between N- and C-terminal domains is similar in both β- and γ-crystallins, the interaction is intermolecular in β-crystallins resulting in oligomers whereas in γ-crystallins the interface is intramolecular which always yields monomers.

The γ-crystallins can be divided into two groups depending on whether they have a high or low temperature, T_c, for phase separation (Blundell *et al.*, 1983; Siezen, Fisch, Slingsby & Benedek, 1985; Summers *et al.*, 1986; Siezen, Wu, Kaplan, Thomson & Benedek, 1988; Broide, Berland, Pande, Ogun & Benedek, 1991). T_c is a measure of protein–protein interactions as a function of temperature. We have previously reported the refinement of the low-T_c protein bovine γB (γII) at 1.9 Å resolution (Wistow *et al.*, 1983) and the high-T_c γE (γIV) at 2.3 Å (White, Driessen, Slingsby, Moss & Lindley, 1989). The low-T_c bovine protein γC (γIIIb) has been reported at 2.5 Å (Chirgadze *et al.*, 1986, 1991). The lattice contacts of the two classes of proteins have been analyzed in order to compare the protein–protein contacts in the crystal state (Sergeev *et al.*, 1988; White *et al.*, 1989).

The original refinement (Wistow *et al.*, 1983) of bovine γB (γII) at 1.9 Å using *RESTRAIN* (Moss & Morfew, 1982) was complicated by uncertainties in the chemically derived sequence (Croft, 1972). A synchrotron X-radiation data set collected at EMBL, Hamburg, was then used to refine the structure along with cDNA sequences derived from rat γ-crystallins (Moormann, den Dunnen, Bloemendal & Schoenmakers, 1982) to a nominal resolution of 1.6 Å (Summers, 1986). The arrangement of surface charges and the disposition of sulfhydryl, aromatic and arginyl side chains were reported based on this refinement (Summers *et al.*, 1984). Since then the homologous bovine sequence has been derived (Bhat & Spector, 1984; Hay, Woods, Church & Petrash, 1987) and has been incorporated into the refinement. The derived coordinates from this 1.6 Å resolution synchrotron data set along with 118 water molecules are those currently in the Brookhaven Data Bank (1GCR).

In this paper we report further refinement of γB using a second improved data set collected at 1.47 Å and in addition, the highly defined structure along with 230 water molecules. The oxidized and reduced positions of two cysteines are resolved. We present an analysis of the secondary structure, thermal parameters and water structure. The structure reveals a number of water molecules that can be considered as contributing an essential part of the protein structure at 293 (1) K.

Table 1. *Summary of the crystallographic details*

The estimated standard deviations for the cell parameters are given in parentheses.

Wavelength used (Å)	0.86
Nominal resolution (Å)	1.47
Unit-cell parameters (Å)	
$a = b$	57.53 (1)
c	97.95 (2)
Number of unique reflections to ∞	26151
% of reflections with $I > 3\sigma(I)$	71
Completeness (%)	91
R_{merge} (%)	8.4

2. Experimental

2.1. *Data collection*

The crystals of bovine γB, space group $P4_12_12$, were prepared in the presence of dithiothreitol as described previously (Carlisle, Lindley, Moss & Slingsby, 1977). The data were collected to 1.47 Å resolution from a single crystal, less than two months old, on beamline X11 at the EMBL outstation at DESY, Hamburg, Germany, using an Enraf–Nonius oscillation camera with flat plate cassettes. Three Kodak no-screen X-ray films were used per cassette interleaved with stainless steel foils of 25 μm thickness. The prismatic crystal, approximate dimensions 0.5 × 0.5 × 1.0 mm, was mounted with the c axis along the camera rotation axis. The wavelength was 0.86 Å, the collimation was 0.3 mm and the crystal-to-film distance was 80 mm. A 2° oscillation range was used and the crystal was translated through the beam roughly every 20°, in order to minimize radiation damage; a total of 66° of data were collected with a still photograph taken every 6° to monitor crystal slippage. These photographs were a significant improvement on the previous set for which a large number of reflections were unrecorded in the resolution range 3.96 to 2.5 Å due to overexposure of many films and a pronounced solvent ring. 118 540 intensity measurements were processed and merged to give 26 151 independent reflections (91% complete) with an R_{merge} of 8.4%. The data were corrected for Lorentz and polarization effects. Measurements of transmission curves indicated a variation of less than 5%; no corrections for absorption were made. A summary of the crystallographic details is given in Table 1.

2.2. *Refinement*

Restrained least-squares refinement of the atomic coordinates and isotropic temperature factors of the protein molecule were performed using *RESTRAIN* (Haneef, Moss, Stanford & Borkakoti, 1985; Driessen *et al.*, 1989) on a Cray-1S computer at the University of London Computer Centre with the final stages being carried out on an in-house Convex C220 computer.

The refinement was carried out in stages as described in Fig. 1, with the summary of the details

Table 2. *Details of the refinement of γB-crystallin*

(*a*) Refinement statistics

Resolution range (Å)	8.0–1.47
R factor (%)	18.1
The data-to-parameter ratio (*Q* value)	
Number of reflections	25959
Number of restraints	3814
Total number of observations	29773
Number of protein atoms	1474
Number of water molecules	230
Number of positional parameters	5121
Overall *U* and *G*	2
Occupancy	2
Isotropic *U* (individual)	1707
Total number of parameters	6832
Q (data-to-parameter) value	4.4

(*b*) Breakdown of the *R* factor according to resolution ranges in the final refinement cycle of 3GCR

Resolution range (Å)	Number of structure amplitudes	*R* factor (%)*
8.00–2.64	4850	17.1
2.64–2.11	4657	16.0
2.11–1.85	4541	16.0
1.85–1.68	4470	20.3
1.68–1.56	4375	25.6
1.56–1.47	3066	28.9
Totals	25959	18.1

* R factor $= (\Sigma|F_o - |F_c||)/\Sigma F_o|$.

given in Table 2. The bovine γB model (Brookhaven code 1GCR) was used as a starting model with 118 water molecules. Three cycles of whole-molecule rigid-body refinement were carried out in the resolution range 2.5 to 8.0 Å to account for the differences in cell dimensions between the two data sets before proceeding with the refinement of the atomic coordinates. There were small translational shifts (approximately 0.15 Å) in the *y* and *z* directions and a rotational shift of less than 0.1°.

At the end of each refinement stage, phases were calculated with *RESTRAIN* using all the reflections in the resolution range 1.47 to 8.0 Å. Electron-density maps were calculated using the *CCP*4 program suite (SERC Daresbury Laboratory, 1979) with coefficients $2|F_o| - |F_c|$ and $|F_o| - |F_c|$. The structure-factor terms were weighted using modified Sim weights (Read, 1986). Model building was carried out using the graphics program *FRODO* (Jones, 1978, as modified by Dr I. J. Tickle) on Evans and Sutherland Picture Systems. During the course of refinement residues Lys 2, Arg 59, Gln 66, Met 90, Arg 153 and Met 160, whose side-chain electron densities were ill-defined, were replaced by alanines to remove model bias. Five well defined residues, Trp 42, Tyr 50, Arg 79, Asp 108 and Leu 155, were also replaced by alanines as a quality test.

The major change in the structure after the first stage of refinement involved an interchange of the carboxy terminus with the tyrosyl side chain (Fig. 2). Other changes were in positions of the side chains, mainly on the surface, such as Arg 9, Gln 26, Ser 30, Arg 36, Arg 47, Gln 51, Arg 59, Gln 66, Gln 67, Met 90, Arg 95, Arg 115, Met 136, Arg 142, Arg 153, Leu 155, Lys 163, Phe 173 and Tyr 174. Coupled second-site occupancies were allocated to Cys 18, Cys 22 and Arg 95 on the basis of the electron density. Starting values for the occupancies were 50% and for the atomic isotropic U_{iso} were 0.2 Å^2. However, halfway

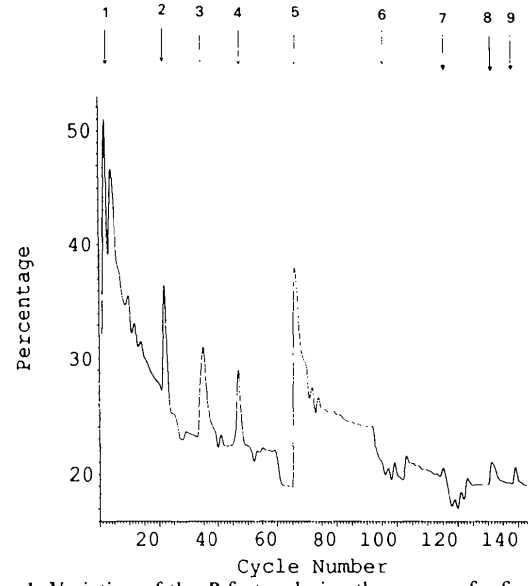


Fig. 1. Variation of the *R* factor during the course of refinement. The refinement was carried out in nine stages, each consisting of a model-building session and restrained least-squares refinement, as indicated by the vertical numbered arrows. The refinement included the Cartesian coordinates (with water molecules) and the isotropic thermal parameter U_{iso}, for each non-hydrogen atom. All non-hydrogen atoms were refined with a fixed occupancy factor of unity, except for the side chains of Cys 18 and Cys 22 (and Arg 95 in the early stages), which were given coupled second-site occupancies. Hydrogen atoms were not included in the refinement since they were not readily locatable in difference Fourier syntheses at 1.47 Å. The initial cycles of each stage were carried out at lower resolution ranges, gradually increasing the upper resolution limit to the limit of 1.47 Å. In the final cycles of each stage, refinement was carried out using all the reflections in the range 1.47 to 8.00 Å. At stage five an omit map was calculated where certain residues (Lys 2, Trp 42, Tyr 50, Arg 59, Gln 66, Arg 79, Met 90, Asp 108, Arg 153, Leu 155 and Met 160) were replaced by alanines in the refinement.

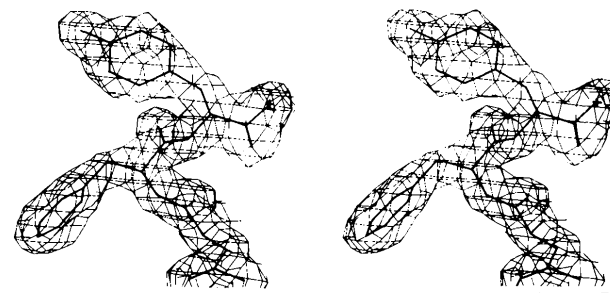


Fig. 2. A stereo plot of the new positions for Phe 173 and Tyr 174 at the carboxy terminus. The electron density is contoured at 0.6 r.m.s.

through the refinement a single occupancy was assigned to Arg 95, the other site being assigned to water molecules. Additional peaks in the difference Fourier maps were assigned to water molecules if situated such that at least one hydrogen-bond contact was made. In the final map 230 water molecules were identified; these were labelled such that water 201 had the lowest U_{iso} value, *i.e.* 0.15 Å^2 and water 430 had the highest value, *i.e.* 0.80 Å^2. A Luzzati plot (Luzzati, 1952) gives an overall estimate of the coordinate accuracy of 0.17 Å.

2.3. *Programs used for analysis*

The solvent-accessible surface area was calculated using a rolling sphere probe, radius 1.4 Å (Lee & Richards, 1971, modified by S. Hubbard). Residues with relative solvent-accessible areas of $\leq 7\%$ were classed as buried. The program *CONTACT* (SERC Daresbury Laboratory, 1979) was used to calculate all the hydrogen bonds made by a given water molecule, including symmetry-related molecules. Criteria for selection included a geometric constraint (angle at the hydrogen $< 120°$ and that at the oxygen $< 90°$ were excluded), as well as the distance restriction and multiple contacts up to six.

3. Results and discussion

3.1. *Temperature factors*

Atomic isotropic mean-square displacement parameters were refined for all atoms including solvent oxygen atoms. The mean U_{iso} value for all 1474 protein atoms (excluding hydrogens) is 0.20 (14) Å^2 while for 696 main-chain atoms the value is 0.15 (6) Å^2. The mean value for the solvent atoms is 0.45 (16) Å^2. The overall mean value for all protein and water atoms is 0.24 (17) Å^2. Plots of the U_{iso} values for the main-chain and the side-chain atoms are shown in Fig. 3. The main-chain plot readily shows, as expected, that the mobile regions of the structure are correlated with the linker region, the C-terminal tail and the long arch between β-strands *c* and *d* particularly in motifs 1 and 3. The side-chain plot demonstrates the mobile nature of the longer functional groups.

3.2. *Torsion angles and secondary structure*

All peptide planes are in the *trans* conformation. A Ramachandran plot (Ramachandran & Sasisekharan, 1968) of the main-chain torsion angles is shown in Fig. 4(*a*). The lower plot defines the seven conformation types used to describe the secondary structure of γB-crystallin (Efimov, 1986; Wilmot & Thornton, 1990; Morris, MacArthur, Hutchinson & Thornton, 1992). Angles typical of the β-extended region (B and P) account for 65.1% of the residues, 23.3% have right-handed α-helical angles (A and a), 6.4% have left-handed α-helical angles (L and G) and 5.2% are in the glycine region (E). The assignment of individual residues as B, P, A, a, L, G and E conformation is shown in Fig. 4(*b*) where the four motifs of γB are aligned with topologically equivalent residues arranged in vertical rows. The alignment also indicates that 42.5% of residues are in the four sets of β-strands (*a*, *b*, *c* and *d*) making up the four β-sheets (Fig. 5). This assignment of β-strands excludes the residues that are in the folded β-hairpin between the *a* and *b* β-strands. This highly conserved, yet irregular, part of the structural motif is critically involved in making tertiary interactions (Lapatto *et al.*, 1991). Although over a quarter of the torsion angles are in α-helical conformation, only in one region, that of the arch connecting the *c* and *d* β-strands of each motif, are they consecutive (Fig. 4*b*). In motifs 1, 2 and 4 they form a single turn of distorted right-handed 3_{10} α-helix and in motif 3 there is one turn of right-handed 3_{14} α-helix. These regions comprise 9.8% of the total residues, or 42% of the residues that have right-handed α-helical angles. Obviously, many of the residues that have α-helical angles are not in helical secondary structure. For example, the topologically conserved glutamates (Glu 7, Glu 46, Glu 94, Glu 135) have right-handed α-helical angles, yet are the end residues of the *a* β-strand in each of the four β-sheets: Asp 38 in the *d* β-strand of motif 1 and its equivalent in motif 3, Leu 127, have α-helical torsion angles yet are β-bulges within the *d* strand of the solvent-facing β-sheets.

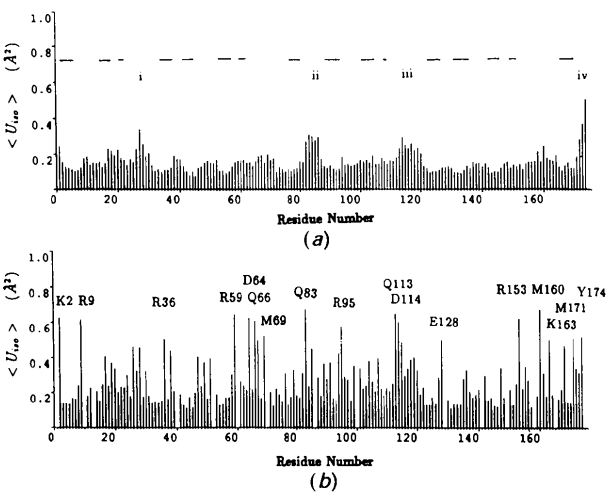
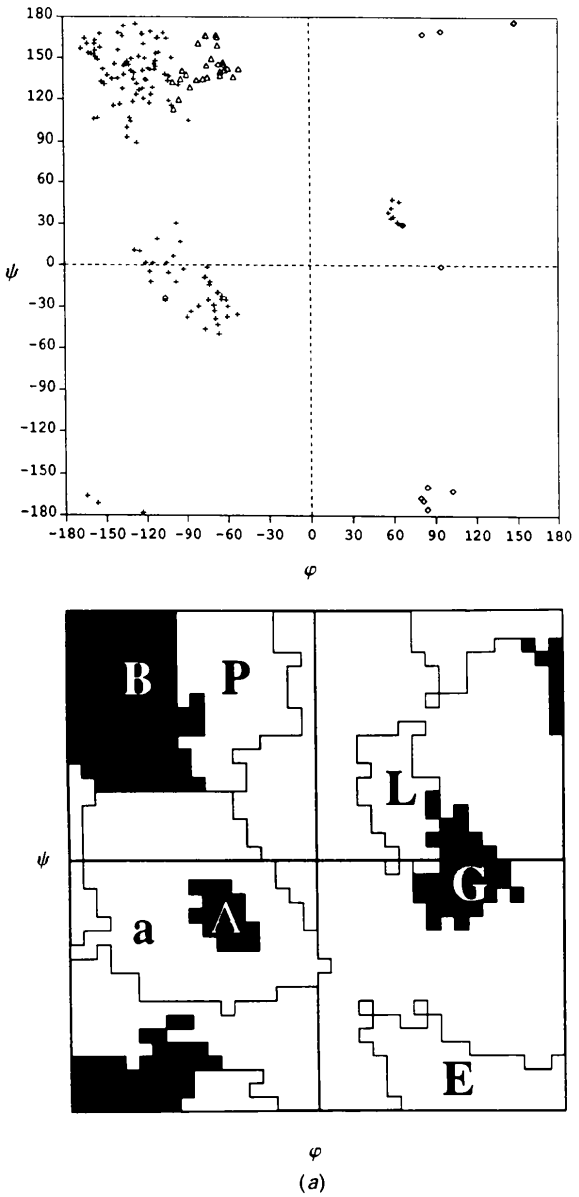


Fig. 3. Plots of average U_{iso} values of γB-crystallin. (*a*) Only the main-chain atoms. The lines above the plot indicate the β strands as defined in Fig. 4(*b*). The four regions with the highest U_{iso} values are (i) the long arch between β strands *c* and *d* of motif 1, (ii) the interdomain connecting peptide, (iii) the long arch between β strands *c* and *d* of motif 3 and (iv) the carboxy-terminus tail. (*b*) Only the side-chain atoms. The residues which have side chains with high U_{iso} values are given by their single amino-acid code and number.

The χ_1 distribution (the rotation around the C^{α}—C^{β} bond) in γB-crystallin is generally trimodal (Bhat, Sasisekharan & Vijayan, 1978; Janin, Wodak, Levitt & Maigret, 1978; McGregor, Islam & Sternberg, 1987). In γB 53% of the residues have χ_1 angles in g^-, 25% in t and 17% in g conformations. 5% of residues have χ_1 angles that fall in between the energetically favourable range: Met 69, Met 90, Glu 17, Glu 94, Asp 64, Asp 97 and Asp 172. Pro, Ala and Gly residues are excluded from this analysis. The χ_2 angles for all glutamates, lysines and arginines in γB are in the *trans* conformation. For aromatic residues χ_2 tends to be $\pm 90^{\circ}$ (Bhat *et al.*, 1978) and this is the case in γB-crystallin, except for Phe 98 in the folded hairpin in motif 3.

3.3. *The sulfur atoms*

γB-Crystallin has an unusually high number of cysteines (7) and methionines (7) in comparison with other molecules (Fahey, Hunt & Windham, 1977). The presence of the sulfur atom and the lack of branching renders the methionine side chain substantially more flexible and polarizable than the branched, and hence comparatively rigid, side chains of leucine and isoleucine (Gellman, 1991). In γB-crystallin Met 43 and Met 90 are buried, whereas the others are on the surface and highly solvent-accessible. The χ torsion angles of the methionine residues of γB-crystallin show a greater variation than other residues, particularly in the χ_3 torsion angle. γ-Crystallins from aquatic lenses tend to have a large number of sulfur-containing residues (Chang, Jiang, Chiou & Chang, 1988). Models of carp γm1 and γm2 built on bovine γB-crystallin have over 14% of their surface covered by methionine residues (Slingsby *et al.*, 1991). The fish lenses have a higher refractive index and are more rigid. The large number of methionines on the surface may allow tighter intermolecular interactions to occur and contribute towards providing a more rigid lens. Protein molecules with high numbers of sulfur atoms may



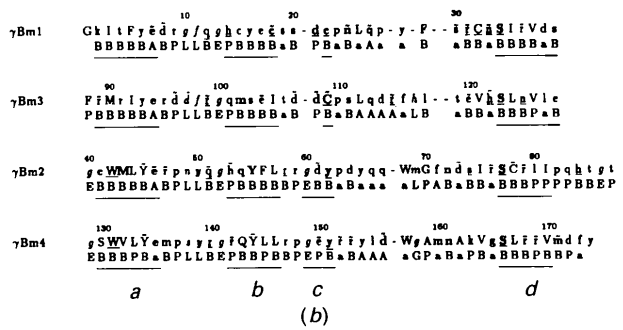


Fig. 4. Analysis of the secondary structure. (*a*) The upper diagram is a Ramachandran plot of the backbone torsion angles of the refined coordinates of γB-crystallin with residues in the poly-prolyl conformation denoted by △, glycine residues are denoted by ◇ and all other residues denoted by +. The lower plot indicates the boundaries used for assigning six types of secondary structure: B (β-extended), P (poly-prolyl), A (right-handed α-helical), L (left-handed α-helical), G and E (regions occupied mainly by glycines) (Wilmot & Thornton, 1990). The less densely populated regions of the right-handed α-helical conformation are denoted by a. Note that many of the γB residues that are assigned to the a region would be classified as γ-conformation by Efimov (1986). (*b*) Sequence alignment of the four motifs of γB-crystallin arranged so that topologically equivalent residues are aligned in vertical rows. The top line is the sequence in single amino-acid notation with the solvent-inaccessible residues in the UPPER CASE, solvent-accessible residues in the lower case; residues with a positive φ in italic; residues whose side chains form a hydrogen bond with another side-chain atom have a tilde, those that form a hydrogen bond to a main-chain amide nitrogen are bold and to main-chain carbonyl are underlined (Overington, Johnson, Sâli & Blundell, 1990). The second line is the secondary structure as defined above. The lines indicate the four β-strands (*a*, *b*, *c* and *d*) of each motif which form part of the β-sheet structure.

contribute to the high refractive index by increasing the polarizability.

The highly refined crystal structure of γB shows that the molecule is partly oxidized with two positions for Cys 22 and two for Cys 18 (Table 3). The electron density indicates that for the major site of Cys 22 the side chain is reduced. The second site is achieved by a rotation of a side-chain torsion angle which results in a disulfide bridge with Cys 18 (Fig. 6). The two sites for Cys 18 are also due mainly to differences in side-chain torsion angles. Both of these sites can presumably be either reduced or oxidized.

The disulfide bridge modelled and refined into the density has an S—S distance of 2.16 Å, compared to the target value of 2.08 Å used. The dihedral energy of this bridge, calculated as described by Katz & Kossiakoff (1986), was $\sim$27.6 kJ mol^{-1} and is high compared with 2.3 kJ mol^{-1} calculated for a left-handed spiral. If the minor site for Cys 18 is built into a disulfide bridge, the resulting left-handed bridge has even higher energy (36.0 kJ mol^{-1}) due to an unfavourable χ_1 dihedral angle (Table 4). The high energy of the disulfide bridge is probably a result of there being only three intervening residues. Our interpretation of the γB structure is that the protein, following reduction with dithiothreitol, comprised seven cysteine residues when originally crystallized. Thereafter, during storage or data collection, cysteines 18 and 22 gradually become oxidized by rotation mainly of the side chain of Cys 22. The disulfide bridge being both high energy and surface accessible would be predicted to be readily reduced.

We have shown that γ-crystallins readily form mixed disulfides with 2-mercaptoethanol and glutathione (Slingsby & Miller, 1985). The mixed disulfides are sufficiently stable to enable their isolation although they can be readily reduced by dithiothreitol. Raman spectroscopy of γB isolated in the absence of added reducing agents showed no signal characteristic of disulfide, whereas following reaction with 2-mercaptoethanol or glutathione there was a drop in the sulfhydryl signal and a corresponding appearance of disulfide signal (Yu, de Nagel & Slingsby, 1989). Solutions of γ-crystallins prepared in the presence of 2-mercaptoethanol show low titres for rapidly reacting sulfhydryl, but they can be recovered following treatment with dithiothreitol (Blundell *et al.*, 1983). However, other workers claim that γB contains a protein–protein intramolecular disulfide bridge that is resistant to reduction with 2-mercaptoethanol (McDermott, Gawinowicz-Kolks, Chiesa & Spector, 1988). The localization of two cysteines in the structure such that an alteration of a side-chain torsion angle of one of them results in a high-energy disulfide, may reflect an unstable transition state in redox reactions with other thiol reagents. Bearing in mind that many of the taxon-specific crystallins are closely related to oxido reductases (Wistow & Piatigorsky, 1988; de Jong,

Table 3. *Occupancies for Cys* 18 *and Cys* 22 *and* U_{iso} *values* (Å^2) *for the* C^β *and* S^γ *atoms*

	Residue	Site one	Site two
Occupancy	Cys 18	61%	39%
U_{iso} (Å^2)	C^β	0.23	0.13
	S^γ	0.25	0.34
Occupancy	Cys 22	58%	42%
U_{iso} (Å^2)	C^β	0.30	0.05
	S^γ	0.18	0.19

Note: The major site for Cys 22 is the sulfhydryl one, where the S^γ atom points away from the Cys 18. The S^γ in the major site of Cys 18 can form a right-handed disulfide bridge and the S^γ in the minor site of Cys 18 can form a left-handed disulfide bridge with the S^γ in the minor site of Cys 22.

Table 4. *Geometries of the two disulfide bridges between Cys* 18 *and Cys* 22 *in γB-crystallin* (3*GCR*)

The directionality of the bridge has been chosen so that the torsion angles from Cys 18 are unprimed and those from Cys 22 are primed. The energies were calculated with the equation obtained from a subroutine of *AMBER* (Weiner *et al.*, 1984) (from Katz & Kossiakoff, 1986). The C^α—C^α distance between Cys 18 and Cys 22 is 7.20 Å. The S—S distance (D) is given in Å, all angles are in ° and the energy in kJ mol^{-1}.

Type	χ_1	χ_2	χ_3	χ'_2	χ'_1	D	Energy
Right-handed	180	99	116	110	−77	2.16	27.61
Left-handed	−112	−101	−116	173	−75	2.09	35.85

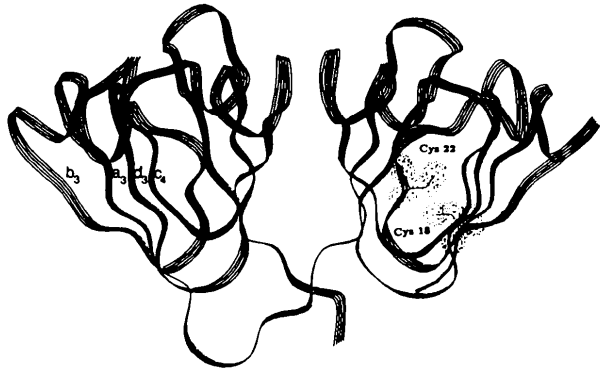


Fig. 5. A ribbon diagram showing the overall fold of γB-crystallin into two similar domains. The two cysteines, 18 and 22, in the reduced state are indicated by dot surface. Cys 18 is on the b_1 strand and Cys 22 on the c_1 of the first Greek key motif (a_1, b_1, c_1 and d_1). Each domain is comprised of two β-sheets. The four β-strands of the solvent-facing sheet of the C-terminal domain are indicated.

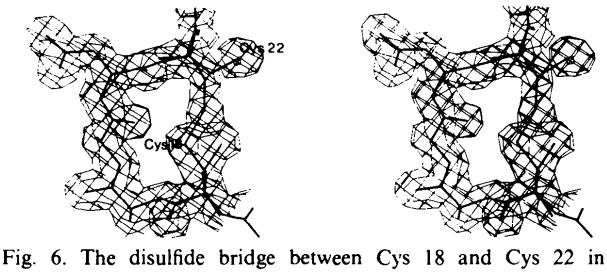


Fig. 6. The disulfide bridge between Cys 18 and Cys 22 in γB-crystallin. Cys 18 and Cys 22 with the sulfur atoms in the minor positions are superimposed on the main model. The electron density is contoured at 0.6 r.m.s.

Hendriks, Mulders & Bloemendal, 1989; Zigler & Rao, 1991) it is possible that γB has a redox role in the lens.

3.4. *The RGD signal peptide*

The tripeptide sequence RGD mediates cell adhesion in many proteins (Pierschbacher & Ruoslahti, 1984). This sequence is present in many γ-crystallins, but not in the related β-crystallins. Comparison of the conformation of this tripeptide in γB-crystallin, α-lytic protease and thermolysin showed that the conformation was quite variable, but that the γ-crystallin recognizes a collagen receptor (Pierschbacher & Ruoslahti, 1987). Solution studies on an active hexapeptide GRGDSP show that it adopts an unusual secondary structure being composed of nested β-bends stabilized by hydrogen bonds between local main-chain atoms and the side chain of the aspartate (Reed *et al.*, 1988). The equivalent sequence in γB, RRGDYP (residues 58–63), has a different structure as the conformation is dominated by hydrogen-bond interactions to a neighbouring β-sheet strand from a distal sequence region that includes an ion pair involving the aspartate (Fig. 7). However, in both structures the critical arginine (59) is in an exposed position free from interactions with its own polypeptide. In γB this arginine has high temperature factors and is interacting with the main-chain carbonyl of His 117, in a symmetry-related molecule. The foot and mouth disease virus VP1 protein has an RGD sequence in a highly disordered loop (Acharya *et al.*, 1989), which forms part of the cell attachment site. In one of the mutant structures of this protein (Parry *et al.*, 1990), the main-chain conformation of this RGD is very similar to the RGD conformation in γB-crystallin (D. Stuart, personal communication). The fact that γB-crystallin RGD sequence has a similar conformation to a protein which has a physiological role in cell attachment would explain receptor binding by γB-crystallin. In summary the backbone region and all side chains of the RRGDYP sequence are well defined by tertiary interactions except the side chain of the arginine which is then presumably free to interact with a receptor.

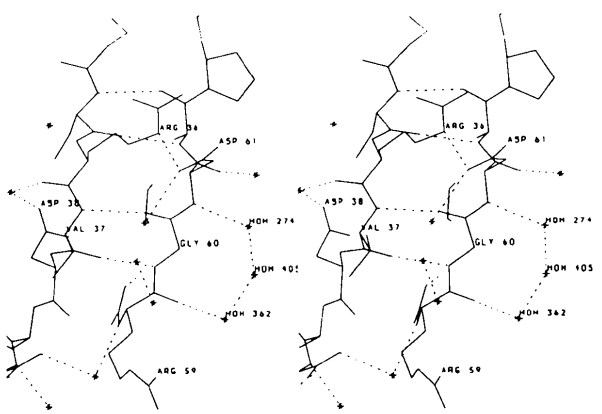


Fig. 7. A stereoview of hydrogen-bonding pattern maintaining the conformation of the RGD sequence which is on the surface of the molecule. The Asp 61 forms an intramolecular ion pair with Arg 36. The sheet structure is extended by a network of water molecules.

Table 5. *Number of waters making a given number of hydrogen bonds* ($2.4 < x < 3.5$ Å) *to protein*

	Number of bonds to protein					
	0 (%)*	1	2	3	4	Total†
Number of water molecules	49 (22)	87	54	30	2	222
Mean U's (Å^2)	0.58	0.48	0.38	0.29	0.22	

* Percentage of waters in the second coordination shell.

† Total of waters selected according to criteria given in text. Note that eight waters make contacts outside the above range.

Table 6. *Analysis of the water hydrogen bonds*

M/C = main chain: S/C = side chain; O = carbonyl oxygen, N = amide nitrogen; $\langle D \rangle$ = average hydrogen bond distance, $\langle U_{H_2O} \rangle$ = average U temperature factor for waters. Note: The average number of contacts to protein per water molecule is 1.3. The analysis was performed using the program *CONTACT* (SERC Daresbury Laboratory, 1979). Limits were $2.4 < D < 3.5$ Å. Angles at the hydrogen atom $< 120°$ and that at the oxygen $< 90°$ were excluded from the analysis. Only the six closest neighbours for each water molecule were included.

Atom	Number	$\langle D \rangle$ (Å)	R.m.s. (Å)	⟨Angle⟩ (°)	$\langle U_{H_2O} \rangle$ (Å^2)
M/C O	102	2.91	0.25	133.3	0.37
M/C N	40	2.92	0.14	162.6	0.32
S/C O	105	2.92	0.31	120.8	0.39
S/C N	46	3.04	0.27	145.7	0.43
HOH	166	3.03	0.31	-	0.37

3.5. *Analysis of the solvent structure*

A total of 230 solvent sites have been identified in the structure. This represents over a third of the expected waters calculated from the crystal density and compares well with other high-resolution structures [*e.g.* actinidin (Baker, 1980)]. 173 of these waters make at least one hydrogen-bonding contact with protein atoms, *i.e.* the water is in the first hydration shell. 49 waters (21% of the total) only make hydrogen bonds with other waters, *i.e.* they are in the second hydration shell (Table 5). A water molecule is considered hydrogen bonded when it is at a distance between 2.4 and 3.5 Å from a protein oxygen or nitrogen or another water molecule. The water molecules have been labelled so that the one with the lowest U temperature factor (0.15 Å^2), has been numbered 201 and the one with the highest (0.80 Å^2) is numbered 430. Thus the first 12 water molecules have U temperature factors similar to main-chain atoms [*i.e.* < 0.15 (6) Å^2] and the next 75 have similar temperature factors to side-chain atoms [*i.e.* < 0.25 (16) Å^2]. Water molecules for which the temperature factor exceeded 0.80 Å^2 and for which there was no electron density at 0.6 r.m.s. in the $2|F_o| - |F_c|$ maps were removed during the course of refinement.

Table 7. *Water molecules making two or more contacts to the protein*

Note: 27 of the first 32 water molecules make multiple contacts to just one protein molecule. # indicates a symmetry-related contact to the adjacent molecule in the unit cell whose location is described by the first three symbols, and the number of the crystallographic symmetry operation, as given in *International Tables for Crystallography* (1983, Vol. A, pp. 358–359), is given by the fourth symbol in parentheses. Underlined water molecules bridge polar atoms between domains of one protein molecule.

Water label	Type of protein atom making contact
Contacts to four protein atoms	
204	O and N Q143, N Q54, NE R142
227	O N33, OD1 N33, OD2 D64, N Y65, HOH
Contacts to three protein atoms + three waters	
216	O and N I75, O D73
226	O E150, OD2 D156, N R152
242	OE2 E7, OG S30, NH1 R31
349	N and O H84, O V170
Contacts to three protein atoms + two waters	
202	O L145, O W157, NE2 Q54
212	N H14, OE2 E7, OD1 D8
220	N S130, O P82, O V170
260	OG S111, N Q113, O G165
Contacts to three protein atoms + one water	
201	OH Y55, O S74, OD1 N72
205	N F88, O T106, NH1 R169
206	NE1 W157, O K163, O N161
209	OE1 E46, ND2 N72, N E#150 (0003)
214	OD1 and OD2 D108, O R169
236	N and OG1 T106, NH1 R169
237	O L80, OD2 D21, N T106
248	OE1 Q12, N G13, OD1 D#114 ($\bar{1}$006)
250	OD1 N125, NH1 R91, OH Y93
272	N and OG S19, O S20
277	O E94, ND1 H122, O E120
338	NH1 and NH2 R76, OD1 #156 (0003)
345	NE R95, NH1 R99, O F173
Contacts to three protein atoms only	
207	OG1 T4, O R36, OD1 D38
208	O F56, OE1 Q143, O W68
234	O G40, OD1 D172, N R59
235	O R147, O P148, OH T151
240	OG1 T4, OE2 E17, NH2 #R140 ($\bar{1}$006)
246	NH1 R89, OE1 E104, OH Y#154 (0$\bar{1}$05)
254	N and O G1, O S19
290	N and O M171, O R169
303	N and OG1 T119, N E120
Contacts to two protein atoms + four water	
258	OD1 N33, OD1 D73
Contacts to two protein atoms + three waters	
211	OH Y45, OE1 Q54
224	N C41, O I81
238	OH Y50, OD2 D#8 ($\bar{1}$0$\bar{1}$4)
263	OD1 D96, O T#106 (0005)
291	OH Y65, O F71
317	ND2 N33, NH2 R#79 (0003)
337	OE1 Q12, O D#61 (0017)
364	OD2 D21, NH2 R79

Water label	Type of protein atom making contact
Contacts to two protein atoms + two waters	
203	O Q54, N Q143
210	ND2 N125, O V126
213	O F29, O I75
217	O G100, N M#102 (0007)
218	OG S130, O P82
228	O K163, OD1 N161
243	O G141, OE1 Q143
252	O M69, OE1 Q143
261	O L118, O V164
273	O G13, OG S#138 ($\bar{1}$006)
275	OE2 E46, OE2 E#150 (0003)
276	N R58, OH Y62
278	O I81, O P82
286	O D156, NH2 R#76 ($\bar{1}$0$\bar{1}$4)
296	O L145, OH Y151
311	OD1 N125, OE2 E150
325	N D73, OH Y65
330	O P110, OG S138
334	NE2 Q83, N D172
Contacts to two protein atoms + one water	
219	OG S30, NH1 R31
221	NE1 W68, O N72
225	N Y154, O E150
230	O C15, O S#138 (1006)
231	N + O S111
245	O G13, O F11
267	O F98, O G100
269	NE R147, O L145
270	OH Y144, N M160
279	NH1 R79, O Y50
316	NE2 N26, NH1 R76
328	OG1 T106, OD1 D107
352	O D61, NE2 Q67
362	O R59, ND1 H#117 (0$\bar{1}$03)
384	O V37, OD1 D61
395	OG S19, O Q#66 ($\bar{1}$006)
398	OG S30, O N49
Contacts to two protein atoms only	
232	OG1 T87, N A#162 (0$\bar{1}$05)
251	OE1 E17, N G#141 ($\bar{1}$006)
255	O I92, O Q101
268	N G40, OD2 D172
281	OG S111, NE2 Q113
282	NE2 Q54, O G52
287	NH1 + NH2 R142
359	OG S20, OE1 Q67
369	O D172, NH1 R58
418	NH1 + NH2 R58

The water hydrogen bonds were analysed with the program *CONTACT* (SERC Daresbury Laboratory, 1979) and the results are given in Table 6. Baker & Hubbard (1984) have defined internal waters as those having three or four hydrogen bonds to the protein. Table 7 includes the 32 waters in γB-crystallin that make three or more hydrogen bonds. Five of these are in the intermolecular region (waters 209, 240, 246, 248 and 338) and seven are in the interdomain region (waters 202, 204, 208, 220, 234, 237 and 349). Nine solvent molecules are very close to the protein (*i.e.* < 2.4 Å). These include three to the aspartate side chains (8, 64 and 73), one to Glu 128, one to Gln 51 and four to arginines (76, 79, 95 and 99). 13 pairs of waters come within 2.4 Å of each other and may represent multiple occupancy sites. The two most deeply buried water molecules (206, 221) hydrogen bond to the indole nitrogens of a pair of topologically equivalent tryptophans (68 and 157), but they connect with other water molecules to the surface, by way of a short solvent channel.

3.5.1. *Main-chain hydrogen bonding and solvation.* A fully extended and solvated γB-crystallin polypeptide of 174 amino-acid residues would be expected to make 511 main-chain to water-molecule hydrogen bonds (the amino and carboxy terminus and eight prolyl nitrogens are excluded). The X-ray structure shows that all of the 165 main-chain N atoms have a hydrogen-bond partner with 67% hydrogen bonding to main-chain carbonyl oxygens, 22% to side chains and 11% to water molecules only (Table 8). Similarly, the main-chain carbonyl oxygens, each capable

Table 8. *Analysis of main-chain hydrogen bonding and solvation*

In this analysis, only one of the bifurcated hydrogen bonds is included. Where such cases occur, a main-chain to main-chain hydrogen bond takes precedent over a main-chain to side-chain hydrogen bond which in turn takes precedent over a main-chain to solvent contact. Where the bifurcated hydrogen bonds are of the same type, the one with the shortest distance and better geometry is considered. For the amide groups, the numbers of waters within 3.5 Å of a nitrogen atom already forming a hydrogen bond with a protein atom are given in parentheses.

Type of hydrogen bond	N-terminal domain	C-terminal domain
(*a*) Main-chain hydrogen-bonding pattern of the nitrogen atoms		
To main-chain carbonyl oxygens	53 (4)	56 (4)
To side chains	19 (8)	19 (8)
To water molecules only	9	9
Numbers excluded*	6	3
(*b*) Main-chain hydrogen-bonding pattern of the oxygen atoms†		
First hydrogen bond and second hydrogen bond		
To main chain and main chain	11	10
To main chain and side chain	10	8
To main chain and water molecule	7	10
To main chain and free	20	25
To side chain and side chain	2	2
To side chain and water molecules	6	4
To side chain and free	4	2
To water molecules and water molecules	9	7
To water molecules and free	9	12
Free and free	9	6

* Amino-terminus and prolyl nitrogens excluded.
† Excludes carboxy terminus.

of making two hydrogen bonds, can make a maximum of 346 hydrogen bonds, out of which 32% hydrogen bond to main-chain nitrogen atoms, 15% to side chains, 23% to water molecules and 30% have no contacts. More than half of the carbonyl oxygens that make only one or no observed hydrogen bonds are buried. The carbonyl oxygens of Cys 32 and Phe 88 and their topological equivalents Val 121 and Gly 129, and Arg 31, do not make a single hydrogen bond. This analysis suggests that these regions could have a tendency to favour local unfolding to gain solvation and that the similarity in the backbone conformation between the two domains extends to the water interactions.

Distributions of water molecules around the main-chain carbonyl and amino groups are consistent with results obtained from analysis of solvation obtained from a collection of high-resolution protein structures (Thanki, Thornton & Goodfellow, 1988). The few water molecules which are outside the clustering form better hydrogen bonds with other polar protein atoms. The average hydrogen-bonding distance is 2.9 Å (Table 6).

The solvent structure plays an important role in domain packing of γB-crystallin. Only three main-chain polar atoms are involved in interdomain hydrogen bonds (N F56—OE1 Q143; N L145—OE1 Q54; and O M171—NH1 R58). 13 water molecules (underlined in Table 7) bridge polar atoms between the two domains (including 12 hydrogen bonds to main-chain nitrogens and/or oxygens in the N-terminal domain and ten to the C-terminal domain). Waters 204, 203 and 252 (together with water 243) form a cluster of highly ordered water molecules which bridge the two domains. Water 220 forms the start of a network of eight water molecules (220—218—422—381—300—365—376—237) which forms a cage shielding the exposed hydrophobic side chain of Pro 82 (Fig. 8). Waters 218, 220 and 237 are interdomain waters making three contacts in the interdomain region and help to stabilize the water network. The other waters are at van der Waals distances away from Pro 82. Fluctuating water cages may form around the interdomain packing hydrophobic side chains (Met 43, Phe 56, Ile 81, Val 132, Leu 145 and Val 170) should the domains oscillate or librate in a less-dense protein phase.

In β-sheet structures, X-ray-observed water molecules occur mainly at the edge of strands, where they extend the sheet structure, or at the ends of strands, where they extend the β-ladder (Thanki, Umrania, Thornton & Goodfellow, 1991). Examples of these are found in γB-crystallin. The water cluster 274—405—362 extends the sheet structure by hydrogen bonding to the short *c* strand of motif two (Fig. 7). A fairly common pattern found in γB-crystallin is that of water molecules bridging the main-chain carbonyl oxygens of residue *i* to *i* + 2 (*e.g.* water molecules 206, 216, 245, 267 and 290, Table 7).

3.5.2. *Water networks.* The 18 observable water molecules that form bridges between symmetry-related molecules and the network of waters in the second hydration shell assist crystal packing. A schematic representation of the major water networks is depicted in Fig. 9. The largest network consists of 35 water molecules: 22 water molecules from one protein are linked together by 13 symmetry-related water molecules. Apart from 202, 218, 220 and 237, which are interdomain water molecules, the rest of the network covers the N-terminal domain surface. This network is extended to a neighbouring N-terminal domain related by a crystallographic two-fold axis. There are five other water networks covering the γB-crystallin surface, predominantly around the N-terminal domain. In the crystal, 33%

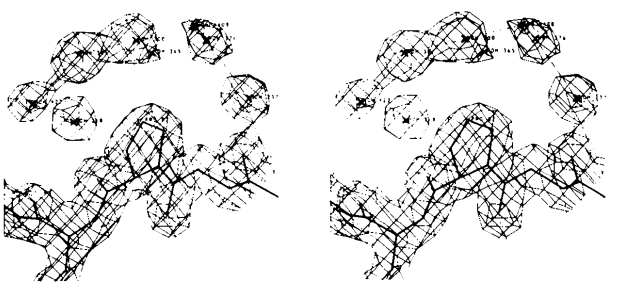


Fig. 8. A stereo picture of the water network shielding the exposed hydrophobic side chain of Pro 82. Note water molecules 300 and 365 are too close (1.95 Å) and are probably the same fluctuating molecule. The electron density is contoured at 0.6 r.m.s.

of the surface (~2800 Å^2) of γB-crystallin is involved in lattice contacts (Sergeev *et al.*, 1988). Most of the ion pairs are on the C-terminal domain and the largest water networks are on the N-terminal domain.

Analysis of the protein hydration of BPTI in solution has led to the conclusion that there are only two different types of water sites, interior or surface, and that the well ordered X-ray-observable surface water molecules have similar residence times for bound water in solution as other surface areas for which no hydration water is seen by X-ray diffraction (Otting, Liepinsh & Wüthrich, 1991). The analysis of the protein hydration in crystals of γB-crystallin indicates a less clear-cut distinction between interior and surface binding sites. There are two deeply buried water molecules (206, 221) that are crucially involved in stabilizing the protein fold. Although they are connected to the bulk water by other water molecules, they may have residency times comparable to those measured for interior hydration sites in BPTI. However, there are 27 observable water sites involving at least three protein water contacts to the surface regions of a single protein molecule. These kinds of water molecules merit an intermediate description.

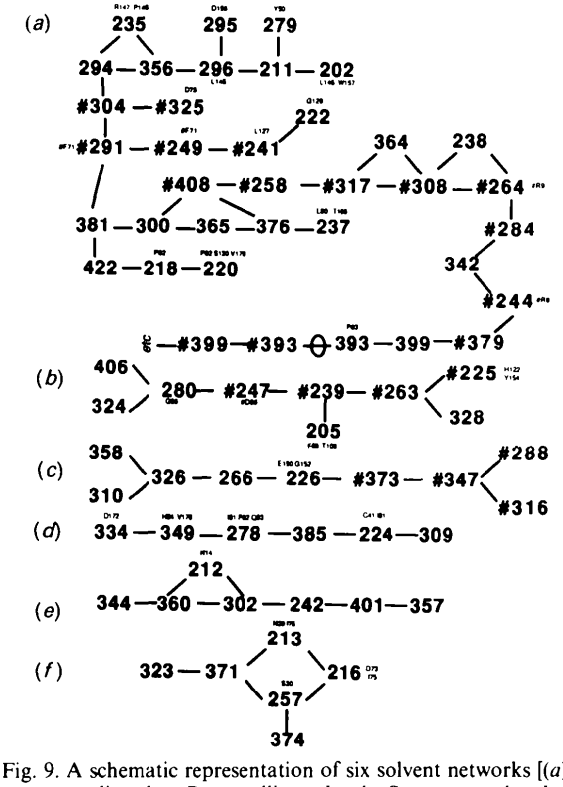


Fig. 9. A schematic representation of six solvent networks [(*a*)–(*f*)] surrounding the γB-crystallin molecule. Symmetry-related water molecules are included (depicted by a #). The residues whose main-chain polar atoms make a hydrogen bond to the water molecule are given by the relevant water molecule.

The protein hydration structure of the γ-crystallins in the crystal state is physiologically relevant as regards protein concentration. The crystals of γB have a protein concentration of 640 mg ml^{-1} (Carlisle *et al.*, 1977), which is intermediate between the total protein concentration of the human lens (~290 mg ml^{-1}) (Fagerholm, Philipson & Lindstrom, 1981) and the rat lens central region (890 mg ml^{-1}) (Philipson, 1969). The refractive index of these lenses is dependent on the protein concentration, being 1.386 for human cortex (Huggart, 1948) and ~1.5 for rat lens core (Philipson, 1969). The protein molecules in the regions of lens with the highest refractive index, although not in a lattice, must be essentially static. This means that protein water sites as determined by X-ray crystallography will have more relevance than a solution estimate where protein molecules have fast correlation times. It is expected that many water molecules, especially those not involving a lattice interaction, will occupy similar sites on the protein in the lens. The general features of the molecular packing in the crystal give an indication of how water molecules can contribute to stabilizing a medium of high refractive index, although in the lens this will be very sensitive to the overall protein concentration, the other crystallin components and a non-periodic array.

In terms of stability, the hydration analysis has indicated those regions of the backbone that are buried yet do not have a hydrogen-bonding partner, invariably the carbonyl group, which may therefore be regions of local instability due to the potential for solvation.*

The support of the Medical Research Council (UK) is gratefully acknowledged. We would also like to thank the EMBL Outstation, c/o DESY Hamburg, for providing synchrotron facilities (Dr Klaus Bartels) and Professor D. Blow and his colleagues for assistance with and use of the film-scanning equipment at Imperial College, London.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 4GCR, R4GCRSF), and are available in machine-readable form from the Protein Data Bank at Brookhaven. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37068 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

References

ACHARYA, R., FRY, E., STUART, D., FOX, G., ROWLANDS, D. & BROWN, F. (1989). *Nature (London)*, **337**, 709–716.

BAKER, E. N. (1980). *J. Mol. Biol.* **141**, 441–484.

BAKER, E. N. & HUBBARD, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.

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. B. (1984). *Human Cataract Formation. Ciba Foundation Symposium*, Vol. 106, pp. 237–247. London: Pitman.

BHAT, S. P. & SPECTOR, A. (1984). *DNA*, **3**, 287–295.

BHAT, T. N., SASISEKHARAN, V. & VIJAYAN, M. (1978). *Int. J. Pept. Protein Res.* **13**, 170–184.

BLUNDELL, T. L., LINDLEY, P. F., MILLER, L., MOSS, D. S., SLINGSBY, C., TICKLE, I. J., TURNELL, W. G. & WISTOW, G. J. (1981). *Nature (London)*, **289**, 771–777.

BLUNDELL, T. L., LINDLEY, P. F., MILLER, L., MOSS, D. S., SLINGSBY, C., TURNELL, W. G. & WISTOW, G. J. (1983). *Lens Res.* **1**, 109–131.

BLUNDELL, T. L., LINDLEY, P. F., MOSS, D. S., SLINGSBY, C., TICKLE, I. J. & TURNELL, W. G. (1978). *Acta Cryst.* B**34**, 3653–3657.

BROIDE, M. L., BERLAND, C. R., PANDE, J., OGUN, O. O. & BENEDEK, G. B. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 5660–5664.

CARLISLE, C. H., LINDLEY, P. F., MOSS, D. S. & SLINGSBY, C. (1977). *J. Mol. Biol.* **110**, 417–419.

CHANG, T., JIANG, Y.-J., CHIOU, S.-H. & CHANG, W.-C. (1988). *Biochem. Biophys. Acta*, **951**, 226–229.

CHIRGADZE, YU. N., NEVSKAYA, N. A., FOMENKOVA, N. P., NIKONOV, S. V., SERGEEV, YU. N., BRAZHNIKOV, E. V., GARBER, M. B., LUNIN, V. Y. & URZUMSTEV, A. P. (1986). *Dokl. Akad. Nauk SSSR*, **209**, 492–495.

CHIRGADZE, YU. N., NEVSKAYA, N. A., VERNOSLOVA, E., NIKONOV, S. V., SERGEEV, YU. N., BRAZHNIKOV, E. V., FOMENKOVA, N. P., LUNIN, V. Y. & URZHUMSTEV, A. (1991). *Exp. Eye Res.* **53**, 295–304.

CHIRGADZE, YU. N., SERGEEV, YU. V., FOMENKOVA, N. P. & ORESHIN, V. D. (1981). *FEBS Lett.* **131**, 81–84.

CROFT, L. R. (1972). *Biochem. J.* **128**, 961–970.

DELAYE, M. & TARDIEU, A. (1983). *Nature (London)*, **302**, 415–417.

DRIESSEN, H., HANEEF, I., HARRIS, G. W., HOWLIN, B., KHAN, G. & MOSS, D. S. (1989). *J. Appl. Cryst.* **22**, 510–516.

EFIMOV, A. V. (1986). *Mol. Biol. (USSR)*, **20**, 250–260.

FAGERHOLM, P. P., PHILIPSON, B. T. & LINDSTROM, B. (1981). *Exp. Eye Res.* **33**, 615–620.

FAHEY, R. C., HUNT, J. S. & WINDHAM, G. C. (1977). *J. Mol. Evol.* **10**, 155–160.

FERNALD, R. D. & WRIGHT, S. E. (1983). *Nature (London)*, **301**, 618–620.

GELLMAN, S. H. (1991). *Biochemistry*, **30**, 6633–6636.

HANEEF, I., MOSS, D. S., STANFORD, M. J. & BORKAKOTI, N. (1985). *Acta Cryst.* A**41**, 426–433.

HAY, R. E., WOODS, W. D., CHURCH, R. L. & PETRASH, J. M. (1987). *Biochem. Biophys. Res. Commun.* **146**, 332–338.

HUGGART, A. (1948). *Acta Ophthalmol. Suppl.* **30**, 1–126.

JANIN, J., WODAK, S., LEVITT, M. & MAIGRET, B. (1978). *J. Mol. Biol.* **125**, 357–386.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

JONG, W. W. DE, HENDRIKS, W., MULDERS, J. W. & BLOEMENDAL, H. (1989). *Trends Biochem. Sci.* **14**, 365–368.

KATZ, B. A. & KOSSIAKOFF, A. (1986). *J. Biol. Chem.* **261**, 15480–15485.

LAPATTO, R., NALINI, V., BAX, B., DRIESSEN, H., LINDLEY, P. F., BLUNDELL, T. L. & SLINGSBY, C. (1991). *J. Mol. Biol.* **222**, 1067–1083.

LEE, B. & RICHARDS, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.

LUBSEN, N. H., AARTS, H. J. M. & SCHOENMAKERS, J. G. G. (1988). *Prog. Biophys. Mol. Biol.* **51**, 47–76.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

MCDERMOTT, M. J., GAWINOWICZ-KOLKS, M. A., CHIESA, R. & SPECTOR, A. (1988). *Arch. Biochem. Biophys.* **262**, 609–619.

MCGREGOR, M. J., ISLAM, S. A. & STERNBERG, M. J. E. (1987). *J. Mol. Biol.* **198**, 295–310.

MOORMANN, R. J. M., DEN DUNNEN, J. T., BLOEMENDAL, H. & SCHOENMAKERS, J. G. G. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 6876–6880.

MORRIS, A. L., MACARTHUR, M. A., HUTCHINSON, E. G. & THORNTON, J. M. (1992). *Proteins Struct. Funct. Genet.* **12**, 345–364.

MOSS, D. S. & MORFEW, A. J. (1982). *Comput. Chem.* **6**, 1–3.

OTTING, G., LIEPINSH, E. & WÜTHRICH, K. (1991). *Science*, **254**, 974–980.

OVERINGTON, J., JOHNSON, M. S., SÂLI, A. & BLUNDELL, T. L. (1990). *Proc. R. Soc. London Ser. B*, **241**, 132–145.

PARRY, N., FOX, G., ROWLANDS, D., BROWN, F., FRY, E., ACHARYA, R., LOGAN, D. & STUART, D. (1990). *Nature (London)*, **347**, 569–572.

PHILIPSON, B. (1969). *Invest. Ophthalmol.* **8**, 258–270.

PIATIGORSKY, J. (1989). *FASEB J.* **3**, 1933–1940.

PIERSCHBACHER, M. D. & RUOSLAHTI, E. (1984). *Nature (London)*, **309**, 30–33.

PIERSCHBACHER, M. D. & RUOSLAHTI, E. (1987). *Science*, **238**, 491–496.

RAMACHANDRAN, G. N. & SASISEKHARAN, V. (1968). *Adv. Protein Chem.* **23**, 283–437.

READ, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

REED, J., HULL, W. E., VON DER LIETH, C.-W., KÜBLER, D., SUHAI, S. & KINZEL, V. (1988). *Eur. J. Biochem.* **178**, 141–154.

SERC Daresbury Laboratory (1979). *CCP4. A Suite of Programs for Protein Crystallography*. SERC Daresbury Laboratory, Warrington, England.

SERGEEV, YU. V., CHIRGADZE, YU. N., MYLVAGANAM, S. E., DRIESSEN, H. P. C., SLINGSBY, C. & BLUNDELL, T. L. (1988). *Proteins Struct. Funct. Genet.* **4**, 137–147.

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.

SLINGSBY, C. (1985). *Trends Biochem. Sci.* **10**, 281–284.

SLINGSBY, C., BAX, B., LAPATTO, R., BATEMAN, O. A., DRIESSEN, H. P. C., LINDLEY, P. F., MOSS, D. S. M., NAJMUDIN, S. & BLUNDELL, T. L. (1991). *Presbyopia Research: from Molecular Biology to Visual Adaptation*, edited by G. OBRECHT & L. W. STARK, pp. 41–48. New York: Plenum Press.

SLINGSBY, C. & MILLER, L. (1985). *Biochem. J.* **230**, 143–150.

SUMMERS, L. J. (1986). PhD Thesis, Univ. of London, England.

SUMMERS, L. J., SLINGSBY, C., BLUNDELL, T. L. B., DEN DUNNEN, J. T., MOORMANN, R. J. M. & SCHOENMAKERS, J. G. G. (1986). *Exp. Eye Res.* **43**, 77–92.

SUMMERS, L. J., WISTOW, G., NAREBOR, M., MOSS, D. S., LINDLEY, P. F., SLINGSBY, C., BLUNDELL, T. L., BARTUNIK, H. & BARTELL, K. (1984). *Pept. Protein Rev.* **3**, 147–168.

THANKI, N., THORNTON, J. M. & GOODFELLOW, J. M. (1988). *J. Mol. Biol.* **202**, 637–651.

THANKI, N., UMRANIA, Y., THORNTON, J. M. & GOODFELLOW, J. M. (1991). *J. Mol. Biol.* **221**, 669–691.

WEINER, S., KOLLMAN, P. A., CASE, D. A., SINGH, U. C., GHIO, C., ALAGONA, G., PROFETA, S. JR & WEINER, P. J. (1984). *J. Am. Chem. Soc.* **106**, 765–784.

WHITE, H. E., DRIESSEN, H. P. C., SLINGSBY, C., MOSS, D. S. & LINDLEY, P. F. (1989). *J. Mol. Biol.* **207**, 217–235.

WILMOT, C. & THORNTON, J. M. (1990). *Protein Eng.* **3**, 479–493.

WISTOW, G. J. & PIATIGORSKY, J. (1988). *Annu. Rev. Biochem.* **57**, 479–504.

WISTOW, G. J., TURNELL, W., SUMMERS, L. J., SLINGSBY, C., MOSS, D. S., MILLER, L. R., LINDLEY, P. F. & BLUNDELL, T. L. (1983). *J. Mol. Biol.* **170**, 175–202.

YU, N.-T., DE NAGEL, D. C. & SLINGSBY, C. (1989). *Exp. Eye Res.* **48**, 399–410.

ZIGLER, J. S. JR & RAO, P. V. (1991). *FASEB J.* **5**, 223–225.