DORNBERGER-SCHIFF, K. (1966). *Acta Cryst.* **21**, 311–322.
DUNITZ, J. D. (1964). *Acta Cryst.* **17**, 1299–1304.
International Tables for X-ray Crystallography (1968). Vol. III. Birmingham: Kynoch Press.
TAKWALE, M. G. & PANT, L. M. (1971). *Acta Cryst.* B**27**, 1152–1158.
UNICS (1967). *The Universal Crystallographic Computing System*(I), edited by T. SAKURAI. Tokyo: The Crystallographic Society of Japan.
WILSON, H. R. (1966). *Diffraction of X-rays by Proteins, Nucleic Acids and Viruses*, p. 28. London: Edward Arnold Ltd.

Acta Cryst. (1978). B**34**, 3653–3657

The Low-Resolution Structure Analysis of the Lens Protein γ-Crystallin

BY T. L. BLUNDELL, P. F. LINDLEY, D. S. MOSS, C. SLINGSBY, I. J. TICKLE AND W. G. TURNELL

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

(*Received* 5 *June* 1978; *accepted* 18 *July* 1978)

The structure of a calf-lens protein, γ-crystallin fraction (II), has been determined at low resolution (5·5 Å) using single-crystal X-ray diffraction methods. Phases for the diffraction data were calculated by the method of multiple isomorphous replacement and the use of anomalous scattering; heavy-atom derivatives were obtained using ethyl mercury thiosalicylate, ethyl mercury chloride, potassium auricyanide and uranyl acetate. The electron density map shows that the molecule is organized as two globular domains each approximately 25 Å in radius. Preliminary crystallographic observations on two other calf γ-crystallins, fractions (III*b*) and (IV), are also reported.

Introduction

The cytoplasm of the fibrous cells of the vertebrate eye lens contains lens specific proteins, the crystallins. In mammals they are heterogeneous and are classified as α-, β- and γ-crystallins; the relative proportions reflect the region of the lens from which they were isolated (see review by Harding & Dilley, 1976). The γ-crystallins are a family of monomeric proteins of molecular weight 20 000 found mainly in the central core region or 'nucleus' of the lens. Amino acid sequence studies have shown that the different γ-crystallin fractions (II), (III*a*), (III*b*), and (IV) are closely related (Croft, 1972, 1973; Slingsby & Croft, 1978). Preliminary crystallographic investigations on fraction (II) have been reported by us (Carlisle, Lindley, Moss & Slingsby, 1977) and on fractions (II) and (III*b*) by Chirgadze, Nikonov, Garber & Reshetnikova (1977).

The tertiary structure of a structural protein should indicate those features of the polypeptide chain folding which enable the protein to have high stability. This is of particular importance to the γ-crystallins as these proteins are synthesized mainly during the embryonic and foetal stages of mammalian development and are retained in the centre of the eye lens with very little protein turnover for the lifespan of the animal. In long-lived species, *e.g.*, man, the central regions frequently become less transparent in old age with the development of senile nuclear cataract. In cataract the proteins of the human lens nucleus gradually denature, leading to polymerization, pigmentation and acquisition of unknown covalent crosslinks other than disulphide bonds (Dilley & Pirie, 1974). Recent chemical analyses clearly indicate the association of cataract with oxidation of protein sulphydryl to protein–protein disulphide bonds (Takemoto & Azari, 1976), this association being particularly striking in the lens-nuclear region (Truscott & Augusteyn, 1977).

One of the aims of the tertiary structure analysis of the bovine γ-crystallin is to further define the environment and reactivity of the six cysteine residues of fraction (II) and to compare these regions with the corresponding regions in the other γ-crystallin fractions. We present here a low-resolution (5·5 Å) three-dimensional structure analysis of γ-crystallin fraction (II) and preliminary investigations of γ-crystallin fractions (III*b*) and (IV).

Experimental

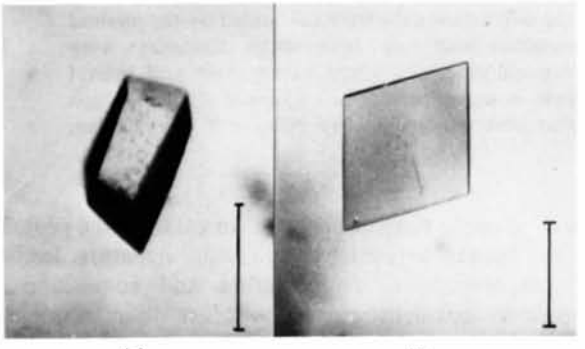
(*a*) *Extraction*

Isolation of calf γ-crystallin and its separation into fractions (I), (II), (III) and (IV) was achieved by methods previously described (Carlisle *et al.*, 1977). A new method of separation of γ-crystallin fraction (III) was devised. To a column of DEAE-cellulose (55 × 2

cm), equilibrated in 0·05 *M* tris-HCl pH 8·5 buffer containing 2-mercaptoethanol, was applied 100 mg of fraction (III) equilibrated in the same buffer. The protein eluted with this buffer is fraction (III*b*). Fraction (III*a*) is then eluted with 0·1 *M* tris-HCl pH 8·5 buffer. γ-Crystallin fraction (IV) was sometimes further purified on DEAE-cellulose by performing chromatography in 0·05 *M* tris-HCl pH 8·5 buffer. The first minor peak is fraction (IV*a*) and the second peak is fraction (IV*b*).

Table 1. *Crystal data on γ-crystallins*

Protein	Space group	Cell dimensions (Å) *a*	*b*	*c*
Fraction (II)	$P4_12_12$	57·8	57·8	98·7
Fraction (III*b*)	$P2_12_12_1$	58·4	69·8	117·7
Fraction (IV)	$C222_1$	35·1	46·2	186·2



(*a*) (*b*)

(*c*)

Fig. 1. (*a*), (*b*) Single crystals of γ-crystallin fraction (IV). The scale bars correspond to 1 mm. (*c*) A 16° precession photograph of the 0*kl* reciprocal lattice layer, γ-crystallin fraction (IV).

(*b*) *Crystallizaton*

Fractions (II) and (III*b*) were crystallized from 6% solutions in sodium phosphate buffer pH 7 at 0°C as described previously (Carlisle *et al.*, 1977). Fraction (IV) was crystallized from a 3% solution using the same phosphate buffer, but at room temperature. The same crystals were obtained using fraction (IV*b*) under the same conditions. Crystal data for the three fractions are given in Table 1. A photograph of the crystals of fraction (IV) and a precession photograph of the 0*kl* reciprocal lattice layer are shown in Fig. 1.

(*c*) *The preparation of heavy-atom derivatives of γ-crystallin fraction* (II)

Most crystals proved to be unstable in the absence of high protein concentrations in the mother liquor. For this reason crystals were first cross-linked overnight in 3% glutaraldehyde in 0·05 *M* sodium phosphate buffer pH 7. The diffraction pattern of the cross-linked crystals showed no changes in cell dimensions and very few changes in intensity when a 20° precession photograph of the *h*0*l* reciprocal lattice layer was compared with the same photograph of the native fraction (II) crystal. Cross-linked crystals were then soaked in 0·05 *M* tris-acetate buffer pH 7 which resulted in no change of diffraction pattern. In view of the several reactive cysteine sulphydryls we began by using mercurials in our search for heavy-atom derivatives. Most mercuric salts resulted in disorder in the crystals, which we thought to be due to a buried or partly occluded sulphydryl. However, reaction with two kinds of mercurials gave useful derivatives. These included the bulky reagent ethyl mercury thiosalicylate (10 m*M* for 18 h) and the small hydrophobic reagents such as ethyl mercuric chloride (13·5 h, saturated solution). It can be seen in retrospect that this was due to a buried sulphydryl being inaccessible to the large mercurial whereas the protein accommodated the small hydrophobic group without too much disruption of the tertiary structure. Changes in intensity were also obtained by soaking the crystals for seven days in 10 m*M* uranyl acetate, and 18 h in potassium auricyanide. All of the heavy-atom solutions were dissolved in 0·05 *M* tris-acetate buffer pH 7, the ethyl mercury chloride requiring traces of ethanol for solubilization.

(*d*) *Data collection and determination of heavy-atom positions*

X-ray intensities for the native and derivative crystals to a maximum Bragg angle $\theta = 13°$ were measured on a Hilger & Watts four-circle diffractometer with Ni-filtered Cu $K\alpha$ radiation. Typically, an ω scan with 1° scan width and 50 s counting time, including background, was performed. The peak

centroid was computed to monitor possible crystal movement and a 'moving window' technique (Tickle, 1975) was used to obtain the net integrated intensity. We measured two equivalents (*hkl* and *khl*) for the native and four (*hkl*, *khl*, *hk̄l*, *kh̄l*) for the derivatives. An empirical absorption correction was applied (North, Phillips & Mathews, 1968). For the derivatives, two measurements of the anomalous difference were obtained for each unique reflection from a pair of Friedel equivalents having the same calculated absorption correction.

The r.m.s. relative changes in intensity for the four derivatives are shown in Table 2. These show no tendency to increase at higher resolution, indicating that the isomorphism of the derivatives is satisfactory (Crick & Magdoff, 1956). The lower estimates of the heavy-atom structure amplitudes, F_{HLE}, were calculated from the protein and derivative structure amplitudes and the anomalous differences (for discussion see Blundell & Johnson, 1976). Harker sections of the Patterson functions at 5·5 Å resolution for $w = c/4$ are shown in Fig. 2. The coefficient used was F^2_{HLE}, except for the uranyl acetate derivative where ΔF^2_{iso} was used (ΔF_{iso} is the isomorphous difference); for this derivative the anomalous differences were not significant and the Friedel equivalents were averaged. The uranyl acetate Patterson could not be interpreted owing to the relatively low degree of heavy-atom substitution. The ethyl mercuric thiosalicylate, potassium auricyanide and ethyl mercuric chloride Pattersons could be interpreted in terms of several major sites.

For each derivative independently, the heavy-atom coordinates, relative occupancies and isotropic temperature factors were refined against F_{HLE} to a maximum resolution of 4·5 Å by a conventional full-matrix structure factor least-squares program. The following procedure was found to give satisfactory results: three cycles of refinement of the overall heavy-atom isotropic temperature factor and the individual occupancies (on an arbitrary scale) were alternated with three cycles for the coordinates and individual isotropic temperature factors. In addition, the relative native/derivative scale factor, which had initially been estimated from a difference Wilson plot, was allowed to vary in every cycle, in order to minimize $\sum_h w_h(F_{HLE} - |F_{H(calc)}|)^2$. Initially unit least-squares weights w_h were used; in the later stages a weighting scheme of the form $w_h \quad (AF^2 + BF + C)^{-1}$ was introduced. The constants A, B and C were chosen to minimize the variation of the mean value of $w_h(F_{HLE} - |F_{H(calc)}|)^2$ taken in ranges of F_{HLE}. Typically $A = 8 \times 10^{-5}$ e^{-2}, $B = -0{\cdot}01$ e^{-1}, $C = 1{\cdot}3$. The $F_{HLE}/F_{H(calc)}$ scale was fixed at unity to ensure that all occupancies were on the same scale.

Table 2. *R.m.s. fractional change in intensity for heavy-atom derivatives* (3·4 Å *resolution data*)

Heavy-atom reagent	$\frac{\langle(\Delta I)^2\rangle^{1/2}}{\langle I\rangle}$ (centric)	$\frac{\langle(\Delta I)^2\rangle^{1/2}}{\langle I\rangle}$ (acentric)
Uranyl acetate	0·41	0·26
Ethyl mercuric thiosalicylate	0·64	0·41
Potassium auricyanide	0·49	0·38
Ethyl mercuric chloride	1·05	0·70

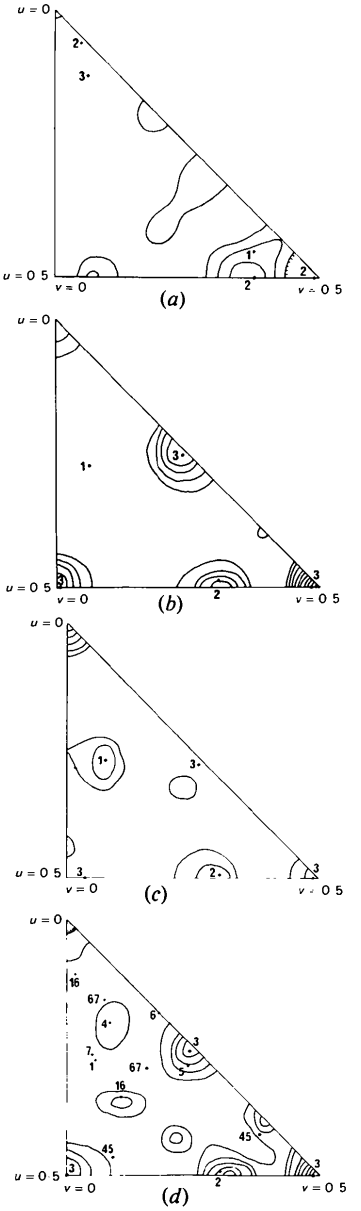


Fig. 2. Heavy-atom Patterson functions at 5·5 Å resolution for (*a*) uranyl acetate, (*b*) ethyl mercuric thiosalicylate, (*c*) potassium auricyanide and (*d*) ethyl mercuric chloride derivatives; Harker section at $w = 0{\cdot}25$. The contour levels are 10, 20, 30 *etc.* for (*a*), (*b*) and (*c*) and 5, 10, 15, 20 *etc.* for (*d*), scaled in each case to an origin peak of 500. The single figures refer to the site numbers in Table 3, and indicate the positions of the self-(Harker-)vectors corresponding to the refined atomic coordinates; the double figures indicate cross-vectors.

Refinement was alternated with computation of an $(F_{HLE} - |F_{H(calc)}|)e^{i\alpha(calc)}$ difference Fourier synthesis weighted by $\min(|F_{H(calc)}|/F_{HLE}, 1)$. The highest significant peak in this map was adopted as a new atomic site and the refinement/difference Fourier procedure iterated, adding one new site at a time. A sensitive test for convergence of the whole process was the disappearance of significant negative difference density at the major site. This test is valid, of course, only if the occupancies and temperature factors have been refined to convergence. Final adjustment of the occupancies as a function of $\sin\theta/\lambda$ was made by refining the $F_{HLE}/F_{H(calc)}$ scale factor in two shells in reciprocal space (16 to 8 Å, and 8 to 5·5 Å). The final difference Fourier synthesis contained no significant positive or negative density.

'Lack of closure' refinement in which the calculated value of the derivative structure factor amplitude given by $F_{PH(calc)} = |F_P + F_H|$ is refined to the observed F_{PH} by varying the heavy-atom positions (Dickerson, Kendrew & Strandberg, 1961) was not used as it depends on an estimate of the protein phase from an independent heavy-atom derivative, but in this case the main derivatives have the same major sites, albeit with different relative occupancies. The choice of space group enantiomorph was achieved in the following way. In a 'cross-difference Fourier' (Dickerson, Kopka, Varnum & Weinzierl, 1967), *e.g.*, combination of the isomorphous difference, ΔF_{iso}, for the auricyanide derivative with protein phases obtained from the ethyl mercuric thiosalicylate derivative in space group $P4_12_12$ (making use of anomalous differences), we observed that both the peak positions and the peak densities were in good agreement with the refined values for the auricyanide derivative. The same calculation in space group $P4_32_12$ gave peak positions and densities which corresponded more closely with the values for the derivative supplying the phases, as would be expected if this were the wrong enantiomorph. The same effect was observed with the ethyl mercuric chloride derivative; in this case we also found that a Fourier synthesis in $P4_12_12$ confirmed the minor sites, whereas one in $P4_32_12$ did not. It should be pointed out that although the cross-difference Fourier syntheses

Table 3. *Heavy-atom parameters and refinement statistics*

Reagent	Uranyl acetate[c]					Ethyl mercuric thiosalicylate				
	x^a	y^a	z^a	Z^{ab}	B^a (Å²)	x	y	z	Z^b	B (Å²)
Heavy-atom sites										
1	0·0856	0·0351	0·1041	0·161	30	0·3959	0·1666	0·1593	0·189	44
2	0·5602	0·0053	−0·0013	0·094	25	0·6046	0·4106	0·0307	0·384	39
3	0·0310	0·5954	0·0529	0·046	8	0·4918	0·2463	0·2485	0·343	41
$R(F_{HLE})$ Overall			–					38·3%		
$R(F_{HLE})$ Centric			57·0%					35·6%		

Reagent	Potassium auricyanide					Ethyl mercuric chloride				
	x	y	z	Z^b	B (Å²)	x	y	z	Z^b	B (Å²)
Heavy-atom sites										
1	0·4032	0·1728	0·1622	0·362	34	0·3915	0·1658	0·1577	0·198	6
2	0·6027	0·4048	0·0345	0·238	42	0·6025	0·4048	0·0319	0·293	9
3	0·4938	0·2682	0·2543	0·113	41	0·4924	0·2496	0·2505	0·256	9
4						0·6445	0·9430	0·0660	0·392	4
5						0·0218	0·2374	0·0718	0·091	18
6						0·8176	0·4999	0·0983	0·111	18
7						0·8426	0·6064	0·1042	0·068	18
$R(F_{HLE})$ Overall			44·4%					42·9%		
$R(F_{HLE})$ Centric			46·1%					47·2%		

Reagent	Uranyl acetate		Ethyl mercuric thiosalicylate		Potassium auricyanide		Ethyl mercuric chloride	
Resolution	16 to 8 Å	8 to 5·5 Å	16 to 8 Å	8 to 5·5 Å	16 to 8 Å	8 to 5·5 Å	16 to 8 Å	8 to 5·5 Å
R.m.s. F_P	302	256	303	258	302	255	302	256
R.m.s. $(F_{PH} - F_P)$	61	42	108	88	91	70	160	133
R.m.s. F_H	43	37	108	90	88	71	130	119
R.m.s. E_H	43	34	35	33	40	37	92	71
R.m.s. F''_H	--	–	9	8	7	6	11	12
R.m.s. E''_H	–	–	18	22	8	11	24	29
Mean $(\alpha_P - \alpha_H)$	89	91	95	91	89	89	90	86

(*a*) Site parameters defining the heavy-atom structure factor $F_H = f_H \Sigma_j^{\text{unit cell}} Z_j \exp(-B \sin^2\theta/\lambda^2)\exp 2\pi i\,(hx_j + ky_j + iz_j)$. (*b*) All site occupancies are on the same, arbitrary scale. (*c*) For this derivative, anomalous scattering was not used, and only centric reflections were used in refinement.

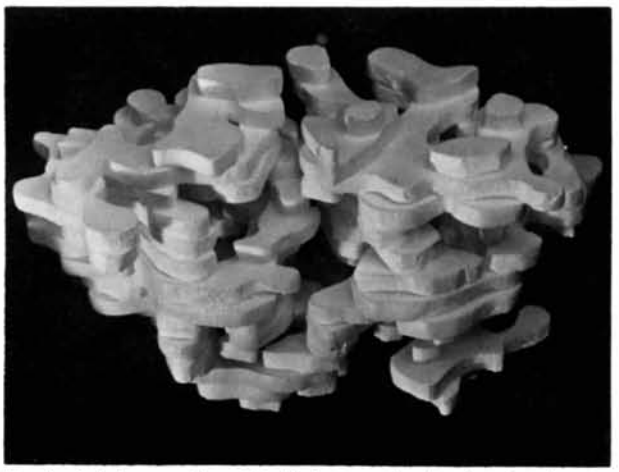


Fig. 3. A balsa-wood model of electron density that is taken to represent one molecule of γ-crystallin fraction (II). For scale see text.

could be used to confirm the results of the refinement, the noise level was about three times that in the ($F_{HLE} - |F_{H(calc)}|$) syntheses and therefore they could not be relied upon to locate minor sites.

A cross-difference Fourier synthesis in $P4_12_12$ with phases calculated from the isomorphous and anomalous data for the three main derivatives gave a site for the uranyl derivative. The refinement and difference Fourier procedure with centric-zone reflections only then generated two further minor sites. Again these were confirmed by the cross-difference Fourier synthesis in $P4_12_12$ but not by one in $P4_32_12$. Finally, phases calculated using the method of multiple isomorphous replacement with all four derivatives in $P4_12_12$ had a mean figure-of-merit of 0·89. The value of the figure-of-merit depends on the estimates of the error E_H in F_H which were obtained from the differences between F_{HLE} and $F_{H(calc)}$ after least-squares refinement of the heavy-atom positions; the error in F''_H cannot be reliably estimated directly [see Blundell & Johnson (1976) for discussion] and was taken as $E''_H = E_H/3$ as suggested by North (1965). Other statistics are shown in Table 3.

(e) *Electron density map*

The best phases and figures of merit (Blow & Crick, 1959) were used to calculate a 5·5 Å resolution electron density map. A balsa-wood model of this is illustrated in Fig. 3.

Results

The electron density map shows a well-defined volume of high density, an ellipsoid with major axes approximately 55 × 30 × 25 Å. This is represented by a balsa-wood model in Fig. 3 which is taken to define the molecular boundary. The model shows that the molecule is organized as two globular domains, each approximately 25 Å in radius. The two domains have large areas of their surfaces in close proximity but at this resolution there is only a single strand of electron density connecting them.

It has been pointed out by Croft (1972) that there is a clustering of four cysteine residues in the *N*-terminal region, between residues 15–32, which is characteristic of fraction (II). Three of the major Hg sites are present in one domain fairly closely together and may represent the positions of cysteines. Further definition of the arrangement of side-chain functions in the γ-crystallin fraction (II) molecule must await the high-resolution study which is under way. It will also be interesting to compare the location of the non-clustered cysteine residues of γ-crystallin fractions (III*b*) and (IV) which are under study in our laboratory.

References

Blow, D. M. & Crick, F. C. (1959). *Acta Cryst.* **12**, 794–802.

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*. London: Academic Press.

Carlisle, C. H., Lindley, P. F., Moss, D. S. & Slingsby, C. (1977). *J. Mol. Biol.* **110**, 417–419.

Chirgadze, Y. N., Nikonov, S. V., Garber, M. B. & Reshetnikova, L. S. (1977). *J. Mol. Biol.* **110**, 619–624.

Crick, F. H C. & Magdoff, B. (1956). *Acta Cryst.* **9**, 901–906.

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

Croft, L. R. (1973). *Ciba Found. Symp.* **19**, 207–226.

Dickerson, R. E., Kendrew, J. C. & Strandberg, B. E. (1961). *Acta Cryst.* **14**, 1188–1193.

Dickerson, R. E., Kopka, M. L., Varnum, J. L. & Weinzierl, J. E. (1967). *Acta Cryst.* **23**, 511–516.

Dilley, K. J. & Pirie, A. (1974). *Exp. Eye Res.* **19**, 59–72.

Harding, J. J. & Dilley, K. J. (1976). *Exp. Eye Res.* **22**, 1–73.

North, A. C. T. (1965). *Acta Cryst.* **18**, 212–215.

North, A. C. T., Phillips, D. C. & Mathews, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

Slingsby, C. & Croft, L. R. (1978). *Exp. Eye Res.* **26**, 291–304.

Takemoto, L. T. & Azari, P. (1976). *Exp. Eye Res.* **23**, 1–7.

Tickle, I. J. (1975). *Acta Cryst.* B**31**, 329–331.

Truscott, R. J. W. & Augusteyn, R. C. (1977). *Exp. Eye Res.* **25**, 139–148.