

of the best presently available refined native enzymes of the subtilisin family and the free cysteine Cys72 suggests, in fact, that it is a member of a subgroup of the subtilisin family (Jany, Lederer & Mayer, 1986; Betzel *et al.*, 1986). It serves as a starting point for crystallographic studies with synthetic dipeptide inhibitors of the chloroketone type (Betzel *et al.*, 1986; Betzel, Wilson, Bellemann, Pal, Bajorath & Saenger, 1988) and work is in progress to solve the structure of the homologous enzyme mesentericopeptidase and thermitase by using molecular replacement, based on the molecular structure of proteinase K (Wilson, 1987, private communication).

We are most grateful to Dr N. Hennrich and Dr H. D. Orth (Merck, Darmstadt), for providing pure samples of proteinase K, and to Dr K. Bartels and Dr H.-D. Bartunik (EMBL/DESY) for providing us with the facilities to collect the synchrotron data. We thank Professor K. D. Jany for making available the amino acid sequence of proteinase K prior to publication and Professor Gassen for communicating changes of the sequence based on the nucleotide sequence of the proteinase K gene. This study has been funded by the German Federal Minister for Research and Technology (BMFT) under contract No. FKZ: B72C050, by Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 9, Teilprojekt A 7), and by Fonds der Chemischen Industrie.

#### References

- ARNDT, U. W. & WONACOTT, A. J. (1977). Editors. *The Rotation Method in Crystallography*. Amsterdam: North-Holland.
- BARTUNIK, H. D., GEHRMANN, T. & ROBRAHN, B. (1984). *J. Appl. Cryst.* **17**, 120.
- BETZEL, CH., PAL, G. P., STRUCK, M.-M., JANY, K. D. & SAENGER, W. (1986). *FEBS Lett.* **197**, 105–110.
- BETZEL, CH., WILSON, K. S., BELLEMANN, M., PAL, G. P., BAJORATH, J. & SAENGER, W. (1988). *Proteins*. Submitted.
- BHAT, T. N. & COHEN, G. H. (1984). *J. Appl. Cryst.* **17**, 244–248.
- COX, E. G. & SHAW, W. F. B. (1930). *Proc. R. Soc. London*, **127**, 71.
- DATTAGUPTA, J. K., FUJIWARA, T., GRISHIN, E. V., LINDNER, K., MANOR, P. C., PIENIAZEK, N. J., SAENGER, W. & SUCK, D. (1975). *J. Mol. Biol.* **97**, 267–271.
- EBELING, W., HENNRICH, N., KLOCHOW, M., METZ, H., ORTH, H. D. & LANG, H. (1974). *Eur. J. Biochem.* **47**, 91–97.
- FOX, G. C. & HOLMES, K. C. (1966). *Acta Cryst.* **20**, 886–891.
- FRENCH, S. & WILSON, K. S. (1978). *Acta Cryst.* **A34**, 517–525.
- HENDRICKSON, W. A. & KONNERT, J. H. (1981). In *Biomolecular Structure, Function, Conformation and Evolution*, edited by R. SRINIVASAN, Vol. 1, pp. 43–57. Oxford: Pergamon Press.
- HENDRICKSON, W. A. & LATTMAN, L. E. (1970). *Acta Cryst.* **B26**, 136–142.
- HIGUCHI, Y., KUSANOKI, M., MATSUURA, Y., YASUOKA, N. & KAKUDO, M. (1984). *J. Mol. Biol.* **172**, 109–139.
- HOLMES, M. A. & MATTHEWS, B. W. (1982). *J. Mol. Biol.* **160**, 623–639.
- JANY, K. D., LEDERER, G. & MAYER, B. (1986). *FEBS Lett.* **199**, 139–144.
- JANY, K. D. & MAYER, B. (1985). *Hoppe-Seyler's Z. Physiol. Chem.* **366**, 485–492.
- JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
- KAHN, R., FOURME, R., GADET, A., JANIN, J., DUMAS, C. & ANDRE, D. (1982). *J. Appl. Cryst.* **15**, 330–337.
- KRAUS, E., KILTZ, H. H. & FEMFERT, U. E. (1976). *Hoppe-Seyler's Z. Physiol. Chem.* **357**, 233–237.
- LUZATTI, P. V. (1952). *Acta Cryst.* **5**, 802–810.
- MACHIN, P. A., WONACOTT, A. J. & MOSS, D. (1983). *Daresbury Lab. News*, **10**, 3–9.
- MACHIN, P. A., WONACOTT, A. J. & MOSS, D. (1984). *Daresbury Lab. News*, **13**, 17–19.
- PÄHLER, A., BANERJEE, A., DATTAGUPTA, J. K., FUJIWARA, T., LINDNER, K., PAL, G. P., SUCK, D., WEBER, G. & SAENGER, W. (1984). *EMBO J.* **3**, 1311–1314.
- PAL, G. P., BETZEL, CH., JANY, K. D. & SAENGER, W. (1986). *FEBS Lett.* **197**, 111–113.
- RAMAKRISHNAN, C. & RAMACHANDRAN, G. N. (1965). *Biophys. J.* **5**, 909–933.
- ROSSMANN, M. G. & ARGOS, P. (1975). *J. Biol. Chem.* **250**, 7525–7532.
- SIM, G. A. (1959). *Acta Cryst.* **12**, 813–815.
- WHITTAKER, E. J. W. (1953). *Acta Cryst.* **6**, 218.
- WIEGERS, U. & HILZ, H. (1981). *Biochem. Biophys. Res. Commun.* **44**, 513–519.
- WILSON, K. & YEATES, D. (1979). *Acta Cryst.* **A38**, 146–157.

*Acta Cryst.* (1988). **B44**, 172–178

## The Use of Pseudosymmetry in the Rotation Function of $\gamma$ IVa-Crystallin

BY H. E. WHITE, H. P. C. DRIESSEN, C. SLINGSBY, D. S. MOSS, W. G. TURNELL AND P. F. LINDLEY

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

(Received 20 April 1987; accepted 28 October 1987)

#### Abstract

Bovine lens  $\gamma$ IVa-crystallin crystallizes in space group C22<sub>1</sub>, with cell dimensions  $a = 35.1$ ,  $b = 46.2$ ,  $c = 186.2$  Å, and contains one molecule in the asymmetric unit. The structure was determined at 3.0 Å

resolution using cross-rotation functions and  $R$ -factor searches with the bovine lens protein  $\gamma$ II-crystallin as the model structure. The rotation function appears to be very sensitive to the resolution range and type of coefficient employed; the use of normalized structure-factor amplitudes gave the best results. The

potential problem of a pseudo solution due to an internal pseudo-twofold axis was put to advantage by aligning this axis parallel to  $z$ . The results of the  $R$ -factor search were well defined. The molecular replacement solution was improved by rigid-body least-squares refinement, initially of the whole molecule, then for the two domains. The  $R$  factor at this stage was 39.4% at 2.3–10.0 Å. The  $\gamma$ IVa structure has an even higher internal symmetry than  $\gamma$ II, since the two domains are related by a rotation around the pseudo-twofold axis of 178.7° as compared with 176.2° for  $\gamma$ II.

### Introduction

Eye lenses contain high concentrations of proteins in order to achieve the refractive index necessary for their function. Lens cells are packed with proteins of different sizes (Bloemendal, 1982; Harding & Crabbe, 1984). The major low-molecular-weight component of these water-soluble crystallins is  $\gamma$ -crystallin, a mixture of monomeric polypeptides of molecular weight  $\sim 20\,000$ , synthesized in early development, and thus occurring predominantly in the central regions of the lens (Björk, 1964; Waley, 1969; Slingsby & Croft, 1973; Slingsby & Miller, 1983).

The  $\gamma$ -crystallins, of which there are five or six, depending on the species, are highly homologous (Slingsby & Croft, 1978; Tomarev, Zinovieva, Chalovka, Krayev, Skryabin & Gause, 1984; Meakin, Breitman & Tsui, 1985; Den Dunnen, Moormann, Lubsen & Schoenmakers, 1986). They are expected to have similar structures because of this and the presence of critical conserved residues (Summers, Slingsby, Blundell, Den Dunnen, Moormann & Schoenmakers, 1986). For bovine  $\gamma$ II and  $\gamma$ IVa this homology is approximately 82%. The structure of bovine  $\gamma$ II has been determined to high resolution (Blundell *et al.*, 1981; Wistow *et al.*, 1983; Summers *et al.*, 1984), and consists of two similar globular domains each comprising two similar 'Greek key' motifs. The two domains pack together with a single connection and are related by a pseudo-twofold axis.

The asymmetric unit of these crystallins therefore contains approximate non-crystallographic symmetry. We describe in this paper the solution analysis of  $\gamma$ IVa by the method of molecular replacement using the structure of  $\gamma$ II, with special attention to the use of this pseudo dyad in the rotation function.

### Experimental procedures

#### Data collection

Bovine lens  $\gamma$ IVa-crystallin crystallizes in space group  $C22_2$ , with cell dimensions  $a = 35.1$  (2),  $b = 46.2$  (2) and  $c = 186.2$  (2) Å (Blundell, Lindley,

Moss, Slingsby, Tickle & Turnell, 1978) with one molecule in the asymmetric unit. The common habit of  $\gamma$ IVa crystals is as very thin flat plates perpendicular to the  $c$  axis (approximately  $0.3 \times 0.3 \times 0.05$  mm), which frequently grow on top of one another with a small misorientation (Slingsby & Miller, 1983). The crystals used for data collection were more rhomboidal ( $0.3 \times 0.3 \times 0.15$  mm).

X-ray intensity data were collected photographically on an Arndt–Wonacott oscillation camera (Arndt & Wonacott, 1977). The crystals were mounted so that  $c$  was parallel to the rotation axis. An initial 44° of data were collected from two crystals at the SERC Synchrotron Radiation Source, Daresbury, UK. A wavelength of 1.488 Å was used with crystal-to-film distances of 68 and 80 mm. The synchrotron was operating at 2 GeV with a beam current of 140 mA decreasing to 119 mA. Another, partially overlapping 74° of data were collected from a third crystal using an Elliott GX6 rotating anode operating at 35 kV and 40 mA. Ni-filtered Cu  $K\alpha$  radiation was used with a crystal-to-film distance of 80 mm. A total of four equivalents ( $hkl$ ,  $h\bar{k}l$ ,  $\bar{h}kl$ ,  $\bar{h}\bar{k}l$ ) was measured at this setting excluding the cusp region, which could not be collected because of the crystal habit. In each case 4° oscillation steps were used, and three films (Ceaverken Reflex 25) were included in each filmpack to give a satisfactory dynamic range.

The photographs were digitized on a Joyce–Loebel Scandig 3 microdensitometer under on-line control of a Data General Nova 3/12 computer. All scanning was performed using a 50  $\mu$ m raster step and an optical density range of 0.0–2.0 units. The digitized data were processed with a program suite based on the Cambridge program *MOSCO* (Nyborg & Wonacott, 1977) as modified by A. J. Wonacott. The inter-filmpack scaling was performed using the scaling algorithm of Fox & Holmes (1966), where a refinable scale and temperature factor are assigned to each film. The scaled data were merged and a residual  $\sum_{hkl} \sum_n |I_{hkl} - \bar{I}_{hkl}| / \sum_{hkl} \sum_n I_{hkl}$  of 10.4% was obtained for 6296 independent reflections to 2.3 Å, constituting approximately 90% of the data at this resolution. No absorption corrections were applied.

#### Rotation function

The 1.6 Å resolution coordinates of the search molecule, bovine  $\gamma$ II-crystallin (Summers *et al.*, 1984), were aligned with the pseudo-twofold axis of the molecule parallel to  $z$ . The rotated molecule was put into an orthogonal cell of  $P1$  symmetry with dimension  $a = b = c = 60$  Å, and structure factors were calculated to 2.1 Å using the atomic temperature factors of  $\gamma$ II with the *FFT* program (Ten Eyck, 1973).

Cross-rotation searches were performed with the program *ALMN*, based on Crowther's (1972) fast

rotation function as modified by E. Dodson. Initial maps were calculated using  $5^\circ$  steps in each of the Eulerian angles  $\alpha$ ,  $\beta$  and  $\gamma$  (30 Bessel functions). In the region of relevant peaks, the step size was then decreased to  $2.5^\circ$  in all angles (60 Bessel functions). The final solution was obtained using  $0.5^\circ$  steps in  $\beta$  and interpolation. The maps were investigated with resolution ranges between 2.3 and 5.0 Å and Patterson radii of 12.0–30.0 Å. An origin cut-off of 3.0 or 6.0 Å was applied. Usually an artificial temperature factor of 20 Å<sup>2</sup> was used. Maps were also calculated using normalized structure-factor amplitudes, using data between 2.3–10.0 and 2.3–20.0 Å.

### R-factor search

The rotational orientation found for  $\gamma$ IVa was applied to the model  $\gamma$ II coordinates. The molecule was put into an orthogonal cell of *P*1 symmetry with dimensions  $a = 35.1$ ,  $b = 46.2$  and  $c = 186.2$  Å, and structure factors were calculated to 3.0 Å using the atomic temperature factors of  $\gamma$ II, with the *FFT* program (Ten Eyck, 1973). Partial structure factors were calculated for each symmetry-related molecule in *C*22<sub>1</sub> (Nixon & North, 1976).

The *R*-factor searches were performed with the program *SEARCH*, written by E. Dodson, using the residual  $\sum |F_o| - |F_c| / \sum |F_o|$  based on the method of Nixon & North (1976). A map was calculated over one asymmetric unit with a resolution range of 5.0–10.0 Å using a 2 Å step size. The unit-cell packing of the 15 peaks with the least residual was examined graphically with the program *FITZ* (Taylor, 1983) on an Evans & Sutherland Picture System II connected to a PDP 11/60 host computer. The region of all peaks which showed acceptable packing was re-investigated using a grid of 0.5 Å.

Possible translation solutions were applied to the rotated coordinates, and the resulting coordinates subjected to ten cycles of a six parameter (three rotational, three translational) rigid-body refinement at 7.0–10.0 Å with the program *CORELS* (Sussman, Holbrook, Church & Kim, 1977) using an overall temperature factor of 20.0 Å<sup>2</sup>. A conventional *R* factor (see above) and a correlation coefficient

$$\frac{n \sum |F_o| F_c - \sum |F_o| \sum F_c}{\{[n \sum |F_o|^2 - (\sum |F_o|)^2][n \sum F_c^2 - (\sum F_c)^2]\}^{1/2}}$$

were calculated at 3.0–10.0 Å for the 3075 reflections with intensities greater than three standard deviations.

### Initial refinement of molecular replacement solution

The molecular replacement solution was refined as one rigid body, initially at 5.0–10.0 Å (633 reflections), then at 3.0–10.0 Å (3075 reflections) using the program *CORELS* with an overall temperature factor

of 20 Å<sup>2</sup>. The two domains of  $\gamma$ IVa were then refined as two rigid bodies at 3.0–10.0 Å (break between residues 85 and 86). Finally two temperature factors per amino-acid residue, one for the main-chain, one for the side-chain atoms, were refined at 2.3–10.0 Å (5959 reflections). In all refinements only those reflections with intensities greater than three standard deviations were used.

## Results

### Determination of the rotational parameters

*The orientation of the search molecule.* Because  $\gamma$ -crystallins have a pseudo-twofold axis between the two domains, which for the  $\gamma$ II backbone is present to an r.m.s. residual of 0.85 Å with a screw component of 0.2 Å for 304 common main-chain atoms, there are expected to be two peaks in the rotation function map, one of which is a pseudo solution with the molecule rotated around the central pseudo dyad. This pseudo-twofold axis was therefore aligned parallel to *z* by rotating the orthonormal coordinates of  $\gamma$ II with the transformation

$$\begin{pmatrix} x_{11r} \\ y_{11r} \\ z_{11r} \end{pmatrix} = \begin{pmatrix} 0.1100 & 0.1966 & -0.9743 \\ -0.8784 & 0.4880 & 0.0000 \\ 0.4755 & 0.8504 & 0.2253 \end{pmatrix} \begin{pmatrix} x_{11} \\ y_{11} \\ z_{11} \end{pmatrix}$$

This alignment has two advantages. First, the solution and pseudo solution will be present in approximately the same  $\beta$ -section with a difference in  $\gamma$  of about  $180^\circ$ , when using the fast rotation function of Crowther, where the first rotation is  $\gamma$  about *z*. This relationship helps to distinguish them from the spurious peaks, facilitating the interpretation. Second, the distinction between true and pseudo solution will be easier, because the distortion of peak shape on one  $\beta$  section is constant.

These two peaks, referred to as *A* and *B*, are indeed found on  $\beta$  sections near  $90^\circ$  with a difference of

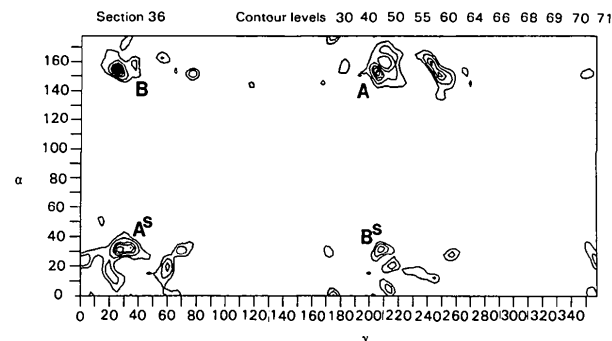


Fig. 1. Cross-rotation map of  $\gamma$ II against  $\gamma$ IVa:  $\alpha$  and  $\gamma$  in  $^\circ$ ,  $\beta = 87.5^\circ$ , resolution 3.0–4.0 Å, Patterson 6.0–22.5 Å, step size  $2.5^\circ$ , *F* coefficients. Peaks *A*<sup>s</sup> and *B*<sup>s</sup> are tails from the symmetry equivalents of peaks *A* and *B* above  $\beta = 90.0^\circ$ .

Table 1. Parameters used to calculate the rotation functions, and the resulting positions and signal-to-noise ratios

For all the calculations the Patterson origin cut-off was 6.0 Å. The artificial temperature factor was 20 Å<sup>2</sup> when the coefficient was the structure-factor amplitude *F*; none was used with normalized structure-factor amplitudes *E*. The signal-to-noise ratio listed for each peak (S/N) is the height of this peak relative to that of the highest one in the map, excluding the pseudosymmetrically related one.

Resolution range (Å)	Patterson radius (Å)	Coefficient	Peak A			S/N	Peak B			S/N	Ratio B/A
			α (°)	β (°)	γ (°)		α (°)	β (°)	γ (°)		
5.0-4.0	15.0	<i>F</i>	—	Not resolved	—	—	Not resolved	—	—	—	
5.0-4.0	22.5	<i>F</i>	175.0	95.0	210.0	1.03	155.0	90.0	25.0	0.93	0.93
4.0-3.0	15.0	<i>F</i>	147.5	90.0	210.0	0.96	152.5	87.5	25.0	0.74	0.77
4.0-3.0	22.5	<i>F</i>	152.5	85.0	205.0	0.93	155.0	87.5	25.0	0.83	0.88
3.0-2.3	15.0	<i>F</i>	155.0	90.0	210.0	0.67	150.0	85.0	25.0	0.91	1.36
3.0-2.3	22.5	<i>F</i>	150.0	90.0	212.5	0.73	157.5	87.5	27.5	1.03	1.36
4.5-3.5	15.0	<i>F</i>	152.5	87.5	210.0	1.10	160.0	90.0	27.5	1.11	1.01
4.5-3.5	22.5	<i>F</i>	155.0	85.0	205.0	0.79	157.5	87.5	25.0	1.01	1.28
5.0-3.0	15.0	<i>F</i>	155.0	90.0	205.0	0.85	160.0	90.0	22.5	1.00	1.18
5.0-3.0	22.5	<i>F</i>	152.5	90.0	205.0	0.84	155.0	85.0	25.0	0.99	1.18
4.0-2.3	15.0	<i>F</i>	147.5	92.5	210.0	1.08	152.5	87.5	25.0	0.92	0.86
4.0-2.3	22.5	<i>F</i>	152.5	85.0	205.0	0.93	155.0	85.0	25.0	0.90	0.88
10.0-2.3	22.5	<i>E</i>	147.5	90.0	212.5	1.00	157.5	87.5	25.0	1.41	1.41
20.0-2.3	22.5	<i>E</i>	147.5	90.0	212.5	0.99	157.5	87.5	25.0	1.44	1.45

approximately 180° in  $\gamma$  between them (Fig. 1). In this cross-rotation study the Patterson of a *P1* cell is rotated over that of a *C222<sub>1</sub>*, giving rise to a rotation function space group *Pbc2<sub>1</sub>* (Moss, 1985). This causes the presence of tails of the symmetry-related peaks *A<sup>s</sup>* and *B<sup>s</sup>* on the same  $\beta$  sections.

#### The choice of the unit cell for the search molecule.

The choice of the unit cell for the search molecule  $\gamma$ II is difficult because of its prolate ellipsoidal shape [57.8 × 27.2 × 25.8 Å, as determined by the method of Taylor, Thornton & Turnell (1983)]. In order to keep the number of intermolecular vectors zero without taking the orientation into account, the dimensions of a *P1* cell should be at least four times the largest semi-axis, *i.e.* 115.5 Å. However, this would constitute a virtually empty cubic cell. One need not search much further than the longest semi-axis, 28.9 Å, and intermolecular vectors larger than this distance will be no problem in practice. A cell of dimensions three times this semi-axis, 86.7 Å, would therefore be more appropriate. However, since the search radius for this aspherical molecule would preferably be the geometrical mean of the ellipsoidal semi-axes (Lifchitz, 1983), 17.1 Å, a further cut in the cell dimensions should be possible, without introducing an unduly large number of intermolecular vectors. Dimensions of 60.0 Å were chosen, which might lead to higher contrast in the rotation function because of the implementation limitations of the rotation function used (1500 strongest reflections only). On the other hand, noise would increase because of intermolecular vectors caused by the dimensions almost being equal to the largest molecular axis, the shortest ones of which can be removed again by using a larger origin cut-off in the Patterson synthesis. Finally, the smaller cell would keep computing costs down, which is an important consideration in view of the fact that many searches have to be done. It is worth noting that using the aligned  $\gamma$ II model in this cell for cross-rotation of  $\gamma$ II against itself the correct solution was readily found (unpublished results).

The effect of the resolution and radius in the rotation function. The heights of the *A* and *B* peaks were investigated at eleven different outer radii for a resolution range of 3.0–4.0 Å. In all cases an origin cut-off of 6.0 Å produced a higher peak than when the inner radius was 3.0 Å (data not shown). At the same time a difference in position was not noticeable for 5° steps. Therefore, for all further studies an origin cut-off of 6.0 Å was used.

Assuming that peaks *A* and *B* represent the pseudo solution and true solution, it was found that the maps are very noisy. The dependence of the signal-to-noise ratio on the resolution range and Patterson radius used are listed for peaks *A* and *B* in Table 1. When using the measured structure-factor amplitudes *F* as coefficients, the noise level was virtually as high as peaks *A* and *B* for all resolution ranges and Patterson radii used. For peak sharpening an artificial temperature factor was used. However, the use of normalized structure-factor amplitudes, *E*, has been found to be advantageous in translation functions (Harada, Lifchitz, Berthou & Jolles, 1981; Tickle, 1985) to produce this effect. When using *E* coefficients in the rotation functions the signal-to-noise ratio was found to increase to 1.4 for peak *B* for data between 2.3 and 20.0 Å, while giving no improvement for peak *A*. A higher resolution cut-off than 20.0 Å was found to reduce the signal-to-noise ratio to approximately one again. Omission of data with *E* coefficients is therefore not advisable. Only with normalized structure-factor amplitudes did one peak, *B*, stand out from the noise.

The relative height of peak *B* to *A* varied, dependent on the resolution shell used. For the highest resolution range of 2.3–3.0 Å (*F* coefficients) and for *E* coefficients, peak *B* was much higher than *A*. When lower resolution data were added, the ratio *B/A* was less favourable, and sometimes completely reversed in value; a consistent trend did not appear. However, within a given resolution shell an increase in the Patterson radius generally led to an increase in the ratio *B/A*. At low resolution

Table 2. Peaks in *R*-factor search maps and their statistics

Peaks are indicated by their origin from the rotation function (*A*1 or *B*), followed by the *R*-factor map number. Where peaks are not completely separated, this is indicated by *a* and *b*. The translation parameters are given in Å. *R* factors at 5.0–10.0 Å have been obtained in the *R*-factor search, while for the other resolution ranges residuals and correlation coefficients were determined with *CORELS*. The r.m.s. shifts have been determined after 10 cycles of rigid-body refinement at 7.0–10.0 Å.

Peak	Translation			Resolution range					R.m.s. shift (Å)
				5.0–10.0 Å		7.0–10.0 Å		3.0–10.0 Å	
	$T_x$ (Å)	$T_y$ (Å)	$T_z$ (Å)	<i>R</i> (%)	<i>R</i> (%)	<i>C</i> (%)	<i>R</i> (%)	<i>C</i> (%)	
<i>A</i> 1-1 <i>a</i>	9.7	5.8	86.8	53.8	49.1	23.9	—	—	0.65
<i>A</i> 1-1 <i>b</i>	9.2	7.7	87.2	54.5	52.4	21.9	—	—	0.75
<i>B</i> -1	13.6	14.4	49.9	45.9	42.5	54.8	48.1	41.5	0.66
<i>B</i> -2 <i>a</i>	12.7	0.0	48.0	52.1	46.9	39.4	54.7	22.6	0.81
<i>B</i> -2 <i>b</i>	13.6	1.0	49.9	51.6	48.0	39.6	53.8	21.8	0.65
<i>B</i> -3 <i>a</i>	13.6	3.4	49.9	52.3	50.9	24.2	53.7	22.3	0.64
<i>B</i> -3 <i>b</i>	13.6	4.3	49.4	52.2	51.7	25.1	56.1	15.5	1.36
<i>B</i> -4	14.0	7.2	49.4	52.5	48.1	41.0	53.6	25.4	0.62
<i>B</i> -5	14.0	8.2	49.4	52.2	46.6	43.6	54.5	21.6	0.58
<i>B</i> -6	14.5	11.0	49.4	51.8	48.0	38.9	53.9	22.7	0.76
<i>B</i> -7	13.2	18.2	49.9	51.9	47.9	39.7	55.9	17.2	0.96
<i>B</i> -8	13.6	20.1	49.4	52.5	45.3	41.2	55.9	16.9	1.13

(4.0–5.0 Å), and using *F* coefficients, for a radius of 15.0 Å, pairs of peaks *B* with the symmetry equivalent of *A*, *A*<sup>s</sup>, and *A* with *B*<sup>s</sup> coalesced (Fig. 1), and an unambiguous determination was not possible.

From the specific pseudosymmetrical relationship of peaks *A* and *B*, and the fact that these peaks often dominate the maps, it is clear that the solution must be contained in them. There are several indications that peak *B* represents the true solution, and peak *A* the pseudo solution. First, peak *B* is the highest at 2.3–3.0 Å with *F* coefficients, a resolution where the difference between the two orientations should be most apparent. This is corroborated by the fact that for *E* coefficients peak *B* is the highest overall. Second, peak *B* is very sharp ( $\sim 4\text{--}6^\circ$  in  $\alpha$ ,  $\gamma$  at height  $\frac{1}{2}$ ), while *A* is broader ( $\sim 5\text{--}10^\circ$ ). Third, and linked to the second argument, for peak *B* the maximum difference in the Eulerian angles as a function of the radius and resolution is small, of the order of 1–2°. For peak *A* this is 3–5°. Nevertheless, it was decided to take both peak *A* and peak *B* into the determination of the translation parameters.

For the final solutions, the resolution range 3.0–4.0 Å was chosen, as found to be suitable in pilot cross-rotation studies of  $\gamma$ II against itself (unpublished results). The Patterson outer radii used were 14.0 and 22.5 Å, and the positions were obtained by interpolation. For peak *B* the difference in the angles at the two radii was negligibly small. We chose the sharper one at 22.5 Å as  $\alpha = 155.5$ ,  $\beta = 86.8$  and  $\gamma = 24.5^\circ$ . For peak *A* the differences were larger. For 14.0 Å the values obtained were  $\alpha = 148.0$ ,  $\beta = 88.0$ ,  $\gamma = 210.8^\circ$  (peak *A*1), and for 22.5 Å  $\alpha = 153.5$ ,  $\beta = 84.8$ ,  $\gamma = 205.5^\circ$  (peak *A*2).

#### Determination of the translational parameters

***R*-factor search.** An initial *R*-factor map was calculated at 5.0–10.0 Å for all rotation-function

solutions, using a coarse 2.0 Å grid to keep computing costs at a reasonable level. The resulting peaks were examined graphically for an acceptable packing, where no rotation axes go through molecules. The surviving peaks were then determined more precisely, using a 0.5 Å grid.

For peak *A*1 from the rotation function, only one with an *R* factor of 55.2% showed reasonable packing. The redetermination with the finer grid showed it to be one peak with two different tops (Table 2), which disappeared completely in the background at 4.0–10.0 Å. For peak *A*2, no peaks which led to acceptable packing and which had *R* factors above the background were produced by the coarse map.

Rotation function peak *B* showed five peaks in the coarse map with acceptable packing (*B*-1 to *B*-5), with *R* factors varying from 50.9 to 52.8%. All these peaks had similar *x* and *z* translations, differing in *y* by up to 15 Å. A fine grid around these five peaks produced three extra closely related peaks (*B*-6 to *B*-8), while peaks *B*-2 and *B*-3 appeared to each have two different tops (Table 2). The *R* factor for the highest peak in the coarse map, *B*-1, decreased from 50.9 to 45.9%, with the next one being *B*-2*b* at 51.6%. When the resolution was increased to 4.5–10.0 Å, all peaks except *B*-1 started to disappear in the background.

On the basis of these results it is clear that rotation function peak *A* does represent the pseudo solution, while peak *B* represents the correct orientation. The ten *R*-factor peaks for *B* fix the *x* and *z* translations at 13.6 and 49.7 Å respectively, making five out of six rigid-body parameters secure. The *y* direction produces considerable noise. However, peak *B*-1 is considered to be the correct solution of the *R*-factor search, because its *R* factor is much lower than the others, and it does not disappear in the background at higher resolution.

**Confirmation of the *R*-factor search solution.** It was necessary to confirm this result by extending the resolution. Especially in *y* the noise was influenced by

the presence of centric reflections. However, since the  $R$ -factor search is computationally very expensive, the  $R$  factors and correlation coefficients at high resolution were determined with *CORELS*. An initial ten-cycle six-parameter rigid-body refinement (three rotational and three translational components) at 7.0–10.0 Å was used to correct for possible gross errors. From the  $R$  factors, peak  $B-1$  was outstanding, with the correlation coefficient being an especially sensitive marker. The correlation coefficients were lower than 24% for the  $A1$  peaks, and for the remaining  $B$  peaks less than 44%. The supposed solution  $B-1$  had 54.8%. Peaks  $A1$  were not considered further. The low resolution ensured a large radius of convergence, with most r.m.s. shifts half a bond distance (Table 2).

Extension of the resolution to 3.0–10.0 Å produced an  $R$  factor of 48.1% for peak  $B-1$  with the next peak being much higher at 53.6% ( $B-4$ ). The correlation coefficient had a considerably better discriminatory power, increasing the signal-to-noise ratio for peak  $B1$  with respect to the next highest one ( $B-4$ ) to 1.63. Peak  $B-1$  showed a good stereochemical packing in the unit cell, even when side chains were taken into account.

#### Rigid-body refinement of the molecular replacement solution

The final transformation to convert the orthonormal coordinates of  $\gamma$ II, with its pseudo-twofold axis aligned along  $z$  to  $\gamma$ IVa was

$$\begin{pmatrix} x_{IV} \\ y_{IV} \\ z_{IV} \end{pmatrix} = \begin{pmatrix} -0.2182 & -0.3563 & -0.9085 \\ -0.3563 & -0.8376 & 0.4141 \\ -0.9085 & 0.4141 & 0.0558 \end{pmatrix} \begin{pmatrix} x_{IIr} \\ y_{IIr} \\ z_{IIr} \end{pmatrix} + \begin{pmatrix} 13.60 \\ 14.37 \\ 49.92 \end{pmatrix}$$

For this cross-rotation solution the six rigid-body parameters were now refined at 5.0–10.0 Å (Fig. 2).

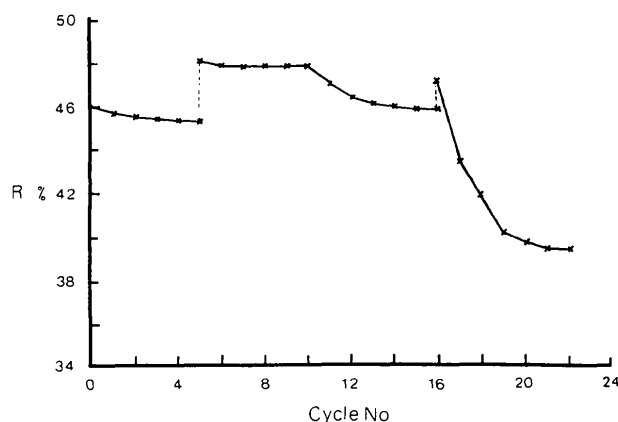


Fig. 2. The progress of least-squares refinement of the molecular replacement solution for  $\gamma$ IVa.  $R$  shows the  $R$  factor as a percentage. In the rigid-body refinement an overall temperature factor of 20.0 Å<sup>2</sup> was used. Cycles 1–5 at 5.0–10.0 Å with one rigid body; cycles 6–10 at 3.0–10.0 Å with one rigid body; cycles 11–16 at 3.0–10.0 Å with two rigid bodies; cycles 17–22 at 2.3–10.0 Å with two temperature factors per residue.

Five cycles reduced the  $R$  factor from 45.9 to 45.3%. A further five cycles at 3.0–10.0 Å gave a final  $R$  factor of 47.9% and correlation coefficient of 43.0%. The total r.m.s. shift with respect to the molecular replacement solution was 0.26 Å, with changes in the angles less than 1° and in translational parameters less than 0.1 Å (note that the convention for the Eulerian angles is different from that of the rotation function).

Because of the presence of two domains in the  $\gamma$ -crystallins, and the fact that a deletion is present in the connecting peptide of  $\gamma$ IVa, compared with  $\gamma$ II, the two domains were refined in six cycles as two independent rigid bodies without restraints. The refinement was completely stable, and the  $R$  factor fell to 45.9%, while the correlation coefficient increased to 48.4%. Inspection of the coordinates showed that the domains had not moved into each other, indicating that they had refined in a correlated way. The total r.m.s. shift with respect to the molecular replacement solution was now 0.88 Å.

Since the overall temperature factor had been held fixed at 20.0 Å<sup>2</sup> up to this stage, the refinement proceeded at 2.3–10.0 Å with two temperature factors per residue, one for the main chain and one for the side chain. The  $R$  factor fell to 39.4% and the correlation coefficient increased to 64.5% in six cycles. The differences between  $\gamma$ II and  $\gamma$ IVa now became apparent in the size of the temperature factors, and were mainly at the expected positions, *i.e.* in most of the side chains where a replacement had taken place in the sequence, in the connecting peptide, where there is a deletion in  $\gamma$ IVa compared with  $\gamma$ II, and in the C-terminal tail, which protrudes out of the molecules in an apparently flexible manner.

#### Discussion

From the rotation-function results it has become clear that although the sequence homology of  $\gamma$ IVa with  $\gamma$ II is about 82%, the signal-to-noise ratio is low for most of the experiments described. The noise level may be explained by several random and systematic errors:

(1) The bilobal structure of the crystallins necessitates a compromise between computing costs and the presence of intermolecular vectors in the choice of a unit cell for the search molecule.

(2) A limitation for the rotation functions has been that the implementation only allowed the use of 1500 reflections and 60 Bessel functions, and therefore a minimum step size of 2.5° in the Eulerian angles  $\alpha$  and  $\gamma$ .

(3) The secondary structure of the  $\gamma$ -crystallins is almost exclusively  $\beta$ -sheet, which may require the use of high-resolution data in the rotation function, because the features of a sheet in a Patterson synthesis are less obvious at intermediate resolution than those of an  $\alpha$ -helix. However, the use of high-resolution data shells

only appears to lead to an increase in the noise, possibly caused in part by the fast rotation function itself. Of course, at high resolution the two molecules would show their differences more readily.

(4) Further there is a global difference between  $\gamma$ II and  $\gamma$ IVa. In  $\gamma$ II the angle for the pseudo-twofold axis is  $176.2^\circ$ . The refinement of the two domains as rigid bodies shows  $\gamma$ IVa to have a more perfect pseudo-twofold axis with an angle of  $178.7^\circ$ , virtually no screw component and a residual of  $0.67 \text{ \AA}$  for main-chain atoms. This would cause the true solution and pseudo solution to be more difficult to distinguish.

(5) The potential problem of the pseudo-twofold axis in the crystallins, giving rise to a systematic pseudo solution, was put to advantage by aligning it parallel to  $z$ , leading to a pinpointing of the correct peak in a sea of noise, and at the same time giving an undistorted comparison of correct and incorrect orientations, by their presence on virtually the same  $\beta$ -section.

The noise level therefore causes the need for 'sharpening' the structure-factor amplitudes. In a crude way this is done by using limited resolution shells and smearing by adding an overall temperature factor to the atomic temperature factors of the search molecule. The use of normalized structure-factor amplitudes appears to circumvent this problem and leads to a more even sampling of reciprocal space. This causes peak  $B$  to rise above the systematic and random noise in relatively few experiments.

In contrast, the  $R$ -factor searches led to a clear result and confirmed the assignment of peak  $B$ . The result of the rotation function was used by Tickle (1985) to examine ways of finding the translation parameters using translation functions. The final test of a molecular replacement solution is always the refinement to high resolution with all data. This procedure, which has currently led to an  $R$  factor of 16% at  $2.3\text{--}5.0 \text{ \AA}$ , will be reported elsewhere.

We thank Dr L. Summers for kindly providing us with the  $1.6 \text{ \AA}$  coordinates of bovine lens  $\gamma$ II. All scanning was performed at Imperial College, University of London, by courtesy of Dr A. Wonacott and Professor D. Blow. We thank Professor T. L. Blundell and Dr I. Tickle for useful discussions. HPCD acknowledges the financial support of the European Molecular Biology Organisation.

## References

- ARNDT, U. W. & WONACOTT, A. J. (1977). Editors. *The Rotation Method in Crystallography*. Amsterdam: North-Holland.
- BJÖRK, I. (1964). *Exp. Eye Res.* **3**, 254–261.
- BLOEMENDAL, H. (1982). *CRC Crit. Rev. Biochem.* **12**, 1–38.
- BLUNDELL, T., LINDLEY, P., MILLER, L., MOSS, D., SLINGSBY, C., TICKLE, I., TURNELL, B. & WISTOW, G. (1981). *Nature (London)*, **289**, 771–777.
- BLUNDELL, T. L., LINDLEY, P. F., MOSS, D. S., SLINGSBY, C., TICKLE, I. & TURNELL, W. G. (1978). *Acta Cryst.* **B34**, 3653–3657.
- CROWTHER, R. A. (1972). In *The Molecular Replacement Method*, edited by M. G. ROSSMAN, pp. 173–178. New York: Gordon & Breach.
- DEN DUNNEN, J. T., MOORMANN, R. J. M., LUBSEN, N. H. & SCHOENMAKERS, J. G. G. (1986). *J. Mol. Biol.* **189**, 37–46.
- FOX, G. C. & HOLMES, K. C. (1966). *Acta Cryst.* **20**, 886–891.
- HARADA, Y., LIFCHITZ, A., BERTHOU, J. & JOLLES, P. (1981). *Acta Cryst.* **A37**, 398–406.
- HARDING, J. J. & CRABBE, M. J. C. (1984). In *The Eye*, 3rd ed., Vol. 1B, edited by H. DAVSON, pp. 207–492. London: Academic Press.
- LIFCHITZ, A. (1983). *Acta Cryst.* **A39**, 130–139.
- MEAKIN, S. O., BREITMAN, M. L. & TSUI, L.-C. (1985). *Mol. Cell. Biol.* **5**, 1408–1414.
- MOSS, D. S. (1985). *Acta Cryst.* **A41**, 470–475.
- NIXON, P. E. & NORTH, A. C. T. (1976). *Acta Cryst.* **A32**, 320–325.
- NYBORG, J. & WONACOTT, A. J. (1977). In *The Rotation Method in Crystallography*, edited by U. W. ARNDT & A. J. WONACOTT, pp. 139–152. Amsterdam: North-Holland.
- SLINGSBY, C. & CROFT, L. R. (1973). *Exp. Eye Res.* **17**, 369–376.
- SLINGSBY, C. & CROFT, L. R. (1978). *Exp. Eye Res.* **26**, 291–304.
- SLINGSBY, C. & MILLER, L. R. (1983). *Exp. Eye Res.* **37**, 517–530.
- SUMMERS, L. J., SLINGSBY, C., BLUNDELL, T. L., DEN DUNNEN, J. T., MOORMANN, R. J. M. & SCHOENMAKERS, J. G. G. (1986). *Exp. Eye Res.* **43**, 77–92.
- SUMMERS, L., WISTOW, G., NAREBOR, M., MOSS, D., LINDLEY, P., SLINGSBY, C., BLUNDELL, T., BARTUNIK, H. & BARTELS, K. (1984). In *Peptide and Protein Reviews*, Vol. 3, edited by M. T. W. HEARN, pp. 147–168. New York: Marcel Dekker.
- SUSSMAN, J. L., HOLBROOK, S. R., CHURCH, G. M. & KIM, S.-H. (1977). *Acta Cryst.* **A33**, 800–804.
- TAYLOR, G. (1983). *J. Mol. Graphics*, **1**, 5–8.
- TAYLOR, W. R., THORNTON, J. M. & TURNELL, W. G. (1983). *J. Mol. Graphics*, **1**, 30–38.
- TEN EYCK, L. F. (1973). *Acta Cryst.* **A29**, 183–191.
- TICKLE, I. J. (1985). In *Molecular Replacement*, edited by P. A. MACHIN, pp. 22–26. Daresbury, UK: Science and Engineering Research Council.
- TOMAREV, S. I., ZINOVIEVA, R. D., CHALOVKA, P., KRAYEV, A. S., SKRYABIN, K. G. & GAUSE, G. G. JR (1984). *Gene*, **27**, 301–308.
- WALEY, S. G. (1969). In *The Eye*, 2nd ed., Vol. 1, edited by H. DAVSON, pp. 299–379. London: Academic Press.
- WISTOW, G., TURNELL, B., SUMMERS, L., SLINGSBY, C., MOSS, D., MILLER, L., LINDLEY, P. & BLUNDELL, T. (1983). *J. Mol. Biol.* **170**, 175–202.