α -Crystallin

FRACTIONATION OF SUBUNITS AND SEQUENCE STUDIES ON AN ISOLATED POLYPEPTIDE

By ROBERT C. AUGUSTEYN AND ABRAHAM SPECTOR

*Department of Ophthalmology, College of Physicians and Surgeons, Columbia University,
New York, N.Y. 10032, U.S.A.*

(Received 17 February 1971)

α -Crystallin was carboxymethylated with radioactive iodoacetic acid in the presence of 7.6 M-urea and then separated into six major fractions by chromatography on DEAE-cellulose in 7 M-urea. Based on the amino acid compositions, specific radioactivities and sodium dodecyl sulphate-gel electrophoresis of the fractions, it was concluded that α -crystallin contains at least four different subunits: DU1A and DU1B, containing no cysteine; a third component represented by DU2B and DU3 containing one cysteine residue per subunit; and DU4, which probably contains two residues of cysteine per subunit. Subunit DU1A was shown to be of sufficient purity for sequence studies. Cyanogen bromide cleavage yielded two peptides, CB-1 and CB-2, in approximately equal amounts as expected. The sum of the molecular weights and amino acid compositions of the peptides were both in excellent agreement with the results obtained for subunit DU1A. The amino acid sequence of the first sixteen residues of peptide CB-1 is: Ser-Leu-Thr-Lys-Asp-Phe-Asp-Glu-Val-Asn-Ile-Asp-Val-Ser-His-Phe-. The sequence of the first seventeen residues of peptide CB-2 is: Asp-Ile-Ala-Ile-Ser-His-Pro-Trp-Ile-Arg-Pro-Ser-Phe-Phe-Glu-Phe-His-. The *N*-terminal sequence of subunit DU1A was shown to be *N*-acetylmethionine followed by peptide CB-2.

α -Crystallin is one of the major structural proteins of the lens. Spector, Li, Augusteyn, Schneider & Freund (1971b) reported that this protein can be divided into three different macromolecular populations on the basis of fractionation by gel filtration. These α -crystallin fractions appear to have identical amino acid compositions and seem to be composed of the same proportion of subunits of 22 500 and 19 500 daltons. The only observed difference between these α -crystallins, apart from their size, is the presence of hexose in the higher-molecular-weight species (Spector *et al.* 1971b). In the present work only the peak 2 material from Agarose 5m columns was used. This α -crystallin fraction represents the lowest-molecular-weight population, in the range 6×10^5 – 9×10^5 daltons. Apart from newly synthesized α -crystallin (Spector, Wandel & Li, 1968), this component represents the narrowest molecular-weight distribution of the different α -crystallins investigated (Spector *et al.* 1971b).

It has been reported that the subunits are held together in the aggregate macromolecule by non-covalent forces and that disulphide bonds are not present (Spector & Katz, 1965). Disaggregation can

be produced by a number of agents such as guanidinium chloride, urea and sodium dodecyl sulphate. However, disulphide formation has been observed in the presence of such disaggregating agents. Therefore, in this study the thiol groups of the protein were carboxymethylated in the presence of 7.6 M-urea before fractionation of the subunits.

There have been numerous attempts to fractionate the subunits of α -crystallin. Early efforts from our laboratory utilizing DEAE-cellulose columns equilibrated in urea suggested a minimum of three subunits (Spector, 1965), whereas Björk (1964), working with a similar system, indicated the possibility of five different subunits. Schoenmakers, Matze, Van Poppel & Bloemendal (1969b) have resolved α -crystallin into two subunit fractions on SE-Sephadex (sulphoethyl-Sephadex) at pH 3 in 7 M-urea. Each fraction was shown to contain at least two components by isoelectric focusing and gel electrophoresis. On the basis of this work, these investigators suggested that α -crystallin is composed of two basic polypeptide chains and two acidic chains. By chromatography on columns of DEAE-cellulose, Palmer & Papaconstantinou (1968) were able to fractionate α -crystallin into three broad

regions. In the present paper, a new method for the fractionation of the subunits from α -crystallin (mol.-wt. 6×10^5 – 9×10^5) on DEAE-cellulose in 7M-urea is presented. Sequence studies on a purified subunit are also reported.

A preliminary report of some of this work was presented by Spector (1970).

MATERIALS AND METHODS

α -Crystallin was prepared by DEAE-cellulose chromatography of a calf lens homogenate followed by gel filtration on Bio-Gel A-5m as described in the preceding paper (Spector *et al.* 1971b). Peak 2 from the Bio-Gel columns was used exclusively for this work. Iodoacetic acid was recrystallized from light petroleum (b.p. 32–50°C) and had m.p. 81.5–82.5°C (literature m.p. 82°C). Iodo[2- 14 C]acetic acid (15.5mCi/mmol) was obtained from Amersham-Searle Corp., Arlington Heights, Ill., U.S.A. Urea solutions were made fresh each day and were stirred for 20 min with Amberlite MB-1 mixed-bed ion-exchange resin (100 g of resin/litre of 8M-urea) immediately before use.

Constant-boiling HCl (b.p. 108°C) was glass distilled three times, bubbled with N_2 and stored in the dark. Sequencer-grade reagents were obtained from Beckman Industries, Spinco Division, Palo Alto, Calif., U.S.A. All other materials were reagent grade and were used without further purification.

Carboxymethylation. Carboxymethylation was performed by a modification of the method of Crestfield, Moore & Stein (1963). α -Crystallin was reduced by dissolving 2 g of the protein in 115 ml of 0.38M-tris buffer, pH 8.6, containing 7.6M-urea, 0.05M-mercaptoethanol, and 1.5mM-EDTA. The solution was bubbled with N_2 and stirred in the dark for 20 min. Then 2 ml of 1M-NaOH containing 0.115mCi of iodo[2- 14 C]acetic acid (15.5mCi/mmol) and 1.08 g of non-radioactive iodoacetic acid, (60 mol of iodoacetic acid/mol of subunit) was added, and the reaction mixture was stirred in the dark for exactly 15 min. The reaction was stopped by the addition of 5 mol of mercaptoethanol/mol of iodoacetic acid, and the reaction mixture was then immediately dialysed against 30 vol. excess of 0.01M-mercaptoethanol at 4°C in the dark. After 1.5 h the dialysis bag was placed in 30 vol. excess of water and dialysis was continued for 18 h with four changes of the water. The dialysate was then freeze-dried and the carboxymethylated α -crystallin was stored at 4°C over a desiccant.

Fractionation of α -crystallin subunits. Carboxymethylated α -crystallin was dissolved in 2mM-phosphate buffer, pH 7.6, containing 7M-urea and 5mM-mercaptoethanol, and was then dialysed overnight against 30 vol. of this buffer. All buffers were prepared from 0.5M- NaH_2PO_4 – K_2HPO_4 , pH 6.80. Approx. 1 g of protein was loaded on a column (2.5 cm $\times$ 60 cm) of DEAE-cellulose equilibrated with the above buffer. The column was operated at room temperature and 10 ml fractions were collected. Protein was eluted with stepwise increases in the concentration of phosphate by using 0.009M-, 0.015M- and 0.08M-phosphate buffers containing 7M-urea and 5mM-mercaptoethanol at pH 7.6. Fractions were assayed by measuring their E_{280} and by determining the radioactivity in 100 μ l

samples. Pooled peaks were dialysed exhaustively against water and were then freeze-dried.

Determination of radioactivity. Radioactivity was determined with a Packard Tri-Carb model 3003 Liquid-Scintillation Spectrometer. Up to 250 μ l of aqueous sample was added to 20 ml of scintillation fluid. The scintillation fluid contained 1400 ml of toluene, 600 ml of ethanol, 8 g of 2,5-diphenyloxazole and 0.2 g of 1,4-bis-(5-phenyloxazol-2-yl)benzene. Specific radioactivities are expressed as c.p.m./ E_{280} unit or as c.p.m./mg, when the relationship between E_{280} and mg of protein was known. This relationship was calculated from the number of E_{280} units taken for amino acid analyses and from the sum of the weights of the amino acids calculated from the amino acid analyses. Good agreement was found with the values obtained in calculations based on the extinctions of the aromatic amino acids and the contents of these amino acids as determined by amino acid analyses.

Amino acid analysis. Measured samples of protein and peptides were hydrolysed under vacuum (0.05 mmHg or less) at 110°C with constant-boiling HCl for 20 and 70 h. The hydrolysates were dried under vacuum and excess of HCl was removed by drying again after the addition of water. Samples were analysed in duplicate with a Beckman-Spinco model 121 automatic amino acid analyser. Reproducibility of $\pm 4\%$ was observed. The analyses were extrapolated to zero time of hydrolysis and the compositions are presented as residues/1000 residues. Yields were based on the recovery of an internal standard, norleucine, which was added to all samples before hydrolysis. Tryptophan was determined spectrophotometrically by the method of Edelhoch (1967) and by *N*-bromosuccinimide oxidation (Spande & Witkop, 1967).

N-Terminal analysis. *N*-Terminal amino acids were determined by the method of Gray & Hartley (1963). DNS-amino acids were identified by t.l.c. by the method of Morse & Horecker (1966).

Cyanogen bromide cleavage. Protein (10 mg/ml) dissolved in 70% (v/v) formic acid was treated at room temperature in the dark for 24 h with a 100-fold molar excess of CNBr. The reaction mixture was then diluted fourfold with water and freeze-dried. The residue was taken up in water and freeze-dried again.

Carboxypeptidase A and B digestion. All digestions were performed at 40°C in 0.25M- NH_4HCO_3 , pH 8.0. Peptides and proteins were digested at a concentration of 2–4 mg/ml with 3 μ M-carboxypeptidase B for 2 h or 10 μ M-carboxypeptidase A for 24 h. At the completion of digestion the samples were freeze-dried and amino acids released were determined by amino acid analysis. Yields were calculated by reference to the recovery of an internal standard, norleucine or histidine, which was added to all samples before digestion.

Sephadex G-50 chromatography. Up to 50 mg of CNBr digest was dissolved in 1 ml of 10% (v/v) acetic acid containing 10% (w/v) sucrose and loaded on a column (1.7 cm $\times$ 150 cm) of Sephadex G-50 equilibrated with 10% (v/v) acetic acid. The column was operated at room temperature at a flow rate of 12 ml/h and 2.0 ml fractions were collected. Peptides were located by measuring the E_{280} of each fraction. Pooled peaks were freeze-dried. All peaks were re-chromatographed under the same condition to minimize cross-contamination.

Gel electrophoresis. Gel electrophoresis in sodium

dodecyl sulphate was performed by the method of Shapiro, Vinuela & Maizel (1967) as modified by Weber & Osborn (1969).

Sequence studies. Sequence studies on 3–4 mg of peptides CB-1 and CB-2 were performed with a Beckman-Spinco model 890 Protein/Peptide Sequencer, with a simple refrigeration system installed on the fraction collector. An atmosphere of N_2 saturated with mercaptoethanol was maintained in the fraction-collector chamber. The sequencer program and the method for conversion of the thiazolines into the phenylthiohydantoin derivatives were essentially the same as those described by Edman & Begg (1967).

Phenylthiohydantoin of amino acids were identified by g.l.c. and by t.l.c. on silica gel. The solvent system V (heptane-propionic acid-ethylene dichloride, 58:17:25, by vol.) of Jeppsson & Sjoquist (1967) was used with untreated silica-gel plates (Eastman Chromagram sheets). Plates were soaked in formamide-acetone (1:3, v/v) for 2 min and then air-dried before solvent systems D (*m*-xylene) and E (formamide-saturated solvent containing 3% propionic acid in water-saturated butyl acetate) of Edman & Sjoquist (1956) were used. Yields of amino acid phenylhydantoin were determined by g.l.c. with a Beckman GC-45 apparatus by using columns (2 mm $\times$ 100 cm) of DC-560 and XE-60 as described by Pisano & Bronzert (1969). With CB-1 the yield of the first residue was assumed to equal 100% and similarly the yield of residue 2 from CB-2 was assumed to be 100%.

RESULTS

α -Crystallin was obtained by DEAE-cellulose chromatography and Bio-Gel A-5m gel filtration as described previously (Spector *et al.* 1971b). A total

of 7 g of peak 2 from the Agarose column was prepared for the work reported in the present paper. This material was carboxymethylated with radioactive iodoacetic acid in 7.6 M-urea to provide a convenient means of determining the thiol content of the subunits as well as following the purification of the fractions and to prevent disulphide formation.

The amino acid composition of the protein before and after carboxymethylation shows that only cysteine has reacted and that there has been no significant change in the content of other amino acids. Five residues of carboxymethylcysteine per 1000 residues were detected. When as much as 2 mg of a hydrolysate of carboxymethylated α -crystallin was loaded on the amino acid analyser column, only trace amounts of cysteine (less than 0.02 mol/mol of protein) were detected. The amino acid composition of carboxymethylated α -crystallin indicated that the protein contained 0.86 mol of carboxymethyl cysteine per 20 500 g. (Since α -crystallin consists of subunits with molecular weights of 19 500 and 22 500 present in a ratio of about 2:1, the average molecular weight of a subunit is therefore approx. 20 500.) Based on amino acid analyses and absorption at 280 nm in 7 M-urea, the $E_{280}^{1\%}$ is 9.9. The specific radioactivity of carboxymethylated α -crystallin was 1263 c.p.m./ E_{280} unit. Since 32 375 c.p.m. represents 1 μ mol of iodo[14 C]-acetic acid, this observed α -crystallin specific radioactivity corresponds to a cysteine content of 0.8 mol/20 500 g. Thus, amino acid analyses and

Fig. 1. Fractionation of the subunits of carboxymethylated α -crystallin on a column (2.5 cm $\times$ 60 cm) of DEAE-cellulose in 7 M-urea at pH 7.6. The column was monitored by measuring the E_{280} (—) and by determining the specific radioactivities (c.p.m./ E_{280} unit, ----) of the fractions. The fractions pooled are shown by the solid bars on the abscissa together with the number assigned to the peak. Arrows indicate points at which the indicated concentrations of phosphate buffer were added.

calculations based on specific radioactivity gave relatively close values for the cysteine content.

For the fractionation of the α -crystallin, 7M-urea was used as the disaggregating agent. Fig. 1 shows a typical elution profile obtained when the disaggregated carboxymethylated α -crystallin was chromatographed on DEAE-cellulose. Over 80% of the E_{280} units loaded on the column was recovered.

The stepwise elution schedule with phosphate buffers containing 7M-urea gave six major fractions. The specific radioactivities (shown as broken lines in Fig. 1) of most of the fractions were relatively constant across the peaks but varied significantly at the trailing edges of the peaks. In all cases except peak 4, only those fractions with a specific radioactivity within a range of $\pm 5\%$ were pooled. The peak 4 material is obviously contaminated, with the variation of specific radioactivity being as much as twofold. The fractions pooled are shown by the solid bars on the abscissa in Fig. 1.

Of the E_{280} units 16% was eluted with the starting buffer as two peaks, DU1A and DU1B. DU2A and DU2B (5% and 20% of the total load respectively) were the major peaks eluted by the 9mM-phosphate buffer. A single peak, DU3, representing 13% of the total E_{280} units was eluted by the 0.15M-phosphate buffer. Another 6% of the E_{280} units was obtained with 0.08M-phosphate buffer. This material was pooled as fraction DU4. Washing with 0.5M-sodium hydroxide eluted only another 1–2% of the E_{280} units which was discarded. The six major fractions accounted for 60% of the E_{280} units loaded on the column. Another 16% was collected as the leading and trailing edges of the peaks, but was not used because of obvious contamination.

The yields and specific radioactivities of the pooled peaks are summarized in Table 1 which shows that fractions DU1A, DU1B and DU2A con-

tain little carboxymethylcysteine. Based on the specific radioactivity, it can be calculated that fractions DU1A and DU1B contain 0.07 and 0.1 residues of carboxymethylcysteine/mol respectively. Such observations indicate that these subunits contain essentially no cysteine and that the carboxymethylcysteine detected is contributed by contaminating polypeptides. A larger contamination was found with fraction DU2A, where 0.3 residue of carboxymethylcysteine/mol was found. Fractions DU2B and DU3 contain approx. 1 residue of carboxymethylcysteine/mol, and fraction DU4, which has about 1.4 residues of carboxymethylcysteine/19 500g, may contain a species with two or more residues of carboxymethylcysteine/mol.

Gel electrophoresis in sodium dodecyl sulphate has revealed that α -crystallin consists of subunits with molecular weights of 19 500 and 22 500 ± 500 in a ratio of about 2:1 (Spector *et al.* 1971b). Trace amounts of components with molecular weights of 12 000–15 000 are also present in the unfractionated material. Estimation of the molecular weights of the subunits revealed that fraction DU1A contains the 22 500-mol.wt. species with trace amounts of low-molecular-weight components (Fig. 2a). Fractions DU2B and DU3 were found to contain the 19 500-mol. wt. species, and the other fractions were found to be mixtures of the 22 500 and 19 500 mol.-wt. and small amounts of low-molecular-weight components. It appears that there are significantly larger amounts of the low-molecular-weight species in some of the fractions from DEAE-cellulose-urea than were present in the carboxymethylated α -crystallin before fractionation. The reason for this result is not understood at the present time.

The amino acid compositions of the fractions were determined and are listed in Table 2. No carboxymethylcysteine was found in fractions DU1A or

Table 1. Summary of the DEAE-cellulose-urea column chromatography of carboxymethylated α -crystallin

Experimental details are given in the text: see also Fig. 1.

Fraction	Yield (%)	Sp. radioactivity		Thiol content (mol/mol of protein)
		(c.p.m./ E_{280} unit)	(c.p.m./mg of protein)*	
DU1A	11	147	103	0.07†
DU1B	5.2	149	128	0.10†
DU2A	5.0	546	490	0.32‡
DU2B	20.2	1810	1660	0.96‡
DU3	12.5	2180	2070	1.17‡
DU4	6.0	2340	2360	1.35‡
Total recovery	75.6			

* The relationship between E_{280} and mg of protein was calculated from the amino acid composition.

† Based on a molecular weight of 22 500.

‡ Based on a molecular weight of 19 500.

Table 2. *Amino acid compositions of fractions from DEAE-cellulose-urea chromatography*

For experimental details see the text. Unless otherwise stated the results are the averages of duplicate 20 h and 70 h hydrolysates extrapolated to zero time of hydrolysis.

Fraction ...	Content (residues/1000 residues)					
	DU1A	DU1B	DU2A	DU2B	DU3	DU4
Amino acid						
Lysine	57	47	49	47	42	40
Histidine	51	43	45	47	43	39
Arginine	86	92	88	82	82	74
Carboxymethyl- cysteine	—	—	2	6	6	8
Aspartic acid	72	74	71	90	93	102
Threonine	36	38	38	30	32	32
Serine	99	100	111	119	120	115
Glutamic acid	104	100	103	106	105	115
Proline	95	85	78	71	67	60
Glycine	49	52	54	57	57	63
Alanine	51	35	42	34	33	34
Valine	52	43	39	49	50	52
Methionine	11	12	12	11	12	11
Isoleucine	52	55	49	50	48	52
Leucine	90	95	97	83	83	85
Tyrosine	12	26	25	31	33	35
Phenylalanine	73	92	85	79	80	78
Tryptophan*	12	12	12	11	12	11

* Determined spectrophotometrically by the method of Edelhoch (1967) and by *N*-bromosuccinimide oxidation (Spande & Witkop, 1967).

DU1B and 0.3 residue was found per mol of fraction DU2A. Approx. 1 mol of carboxymethylcysteine was found/mol of fractions DU2B and DU3. Fraction DU4 contained 1.4 residues of carboxymethylcysteine/mol. These values are in good agreement with the values calculated from the specific radioactivities. The amino acid compositions of the fractions reflect their increasing acidity, as would be expected from their order of elution from the DEAE-cellulose-urea column.

The fractions containing no carboxymethylcysteine, DU1A and DU1B, appear to be the most basic fractions. Fractions DU1A and DU1B are clearly different species as shown by the differences in their amino acid compositions, notably in their contents of lysine, histidine, proline, alanine, valine, phenylalanine and tyrosine.

Since fraction DU2A contains only 0.3 mol of carboxymethylcysteine/mol, it must contain at least two components, one with and one without carboxymethylcysteine. This fraction may include a species that was not found in any of the other fractions, but it will have to be purified further before this can be determined.

The compositions of fractions DU2B and DU3 are very similar, differing significantly only in their contents of lysine, histidine and proline. These polypeptides must represent at least one additional

subunit species since their compositions differ markedly from those of the more basic subunits and since they contain one residue of carboxymethylcysteine/mol.

The amino acid composition of fraction DU4 is somewhat similar to that of fraction DU3, differing in content of carboxymethylcysteine, as well as that of arginine, aspartic acid, glutamic acid, proline and glycine. Since fraction DU4 contains more than one cysteine residue/mol it must represent at least one more subunit species of α -crystallin. However, it is obvious from the carboxymethylcysteine content that this fraction is impure, and with further purification its amino acid composition will probably differ more markedly from that of fraction DU3.

Thus, from a consideration of the specific radioactivities and amino acid compositions of the fractions, it may be deduced that α -crystallin contains at least four different subunits, as represented by fractions DU1A, DU1B, DU4 and DU2B and DU3, which appear to be one entity. Since a number of these subunits are impure, it is possible that there may be a larger number of components.

The specific radioactivity of fraction DU1A indicates that at least 93% of the material represents a component containing no carboxymethylcysteine. Attempts to decrease the radioactivity by re-chromatography under the same conditions

on DEAE-cellulose or on SE-Sephadex in 7M-urea at pH 3.2 and pH 4.6 were unsuccessful. Carboxypeptidase B digestion released 1.08 residues of lysine/mol of fraction DU1A, whereas serine was found to be the major C-terminal amino acid in α -crystallin. Thus, on the basis of specific radioactivity, end-group analysis and gel electrophoresis in sodium dodecyl sulphate, fraction DU1A was considered of sufficient purity for sequence studies.

It would be anticipated from work done in our laboratory and in other laboratories (Hoenders & Bloemendal, 1967; Mok & Waley, 1968) that all α -crystallin subunits would have the N-terminal sequence N-acetylMet-Asp-... Since fraction DU1A contains only two residues of methionine/mol, and one of these would be expected to be in the N-terminal sequence, cyanogen bromide cleavage of fraction DU1A should produce two peptides plus N-acetylhomoserine (or its lactone). As expected, no N-terminal residue could be detected for fraction DU1A. Dansylation of the cyanogen bromide digest of fraction DU1A indicated that there were two major N-terminal residues, serine and aspartic acid (or asparagine) as well as trace amounts of what appeared to be alanine, leucine and glycine. A one-step Edman degradation on the cyanogen bromide digest removed only the phenylthiohydantoins of serine and aspartic acid. Thus, it may be concluded that serine and aspartic acid are the only significant N-terminal amino acids produced by cyanogen bromide cleavage of fraction DU1A. This shows another difference between fraction DU1A and most of the remaining subunits of α -crystallin, since leucine and aspartic acid were the major N-terminal amino acids detected in the cyanogen bromide digest of α -crystallin.

Gel electrophoresis of the cyanogen bromide digest in sodium dodecyl sulphate (Fig. 2b) indicated that fraction DU1A had almost disappeared and was replaced by bands with molecular weights of 12 600 and 9000 (Fig. 2). (The reference polypeptide chains were the same as those reported by Spector *et al.* (1971b). Cytochrome *c* was also used occasionally as a further reference.) The sum of these molecular weights agrees well with the molecular weight of fraction DU1A.

The elution profile of the cyanogen bromide digest of fraction DU1A on Sephadex G-50 is shown in Fig. 3. Fraction DU1A would be expected to emerge with the void volume of the column. The first peak (peak 0) was found in this position, but was only partly resolved from the second component (peak 1). The major fraction (peak 2) was eluted about 30 ml later and was followed by two very minor components.

The recovery of the peptides is shown in Table 3. About 75% of the E_{280} units placed on the column was recovered in this experiment. In later runs,

Fig. 2. Gel electrophoresis in sodium dodecyl sulphate of (a), subunit DU1A; (b), CNBr digest of subunit DU1A; (c), peak 1 and (d), peak 2 from the Sephadex G-50 fractionation of the CNBr digest of subunit DU1A. Experimental details are given in the text. The upper band in each gel is pepsin, used as a standard marker.

Fig. 3. Gel filtration of the CNBr digest of subunit DU1A on a column (1.7 cm $\times$ 150 cm) of Sephadex G-50 (superfine) equilibrated with 10% (v/v) acetic acid. Up to 50 mg of digest in 1 ml of 10% acetic acid containing 10% sucrose was loaded on the column. Peptides were located by measuring the E_{280} of the fractions. Fractions pooled are shown by the solid bars on the abscissa together with the number assigned to the peak.

where larger quantities of material were fractionated, 90% of the starting material was recovered. The E_{280} reflects the aromatic amino acid content of the peptides and not necessarily the amount of material. Therefore, the molar extinction E_{280}^x was determined by correlation of the E_{280} and amino

Table 3. Recovery of cyanogen bromide peptides of subunit DU1A from Sephadex G-50 chromatography

The molecular weights and molar extinctions at 280 nm (E_{280}^M) used in these calculations are based on the amino acid compositions of the purified peptides (see Table 4).

Fraction	Recovery of E_{280} units (%)	$10^{-3} \times E_{280}^M$	Recovery	
			(mg)	(μ mol)
0	8.7	15.5	2.3	0.10
1	7.0	1.38	11.6	0.91
2	50.0	12.9	6.6	0.71
Remainder	8.8	—	—	—
Total recovery	74.5	—	—	—
Initial load	—	—	26.1	1.18

acid compositions of the peptides recovered after a second gel filtration. It was found that the E_{280} of peak 2 material was almost ten times that of peak 1. Calculations of the yields of peptides indicated that peak 1 material was actually a major component of the cyanogen bromide digest. A second gel filtration of peak 1 material resulted in a further removal of peak 0 and indicated that approx. 18% of the E_{280} units, or less than 2 mg% was due to this component. The peptides recovered from the second gel filtration of peaks 1 and 2 materials will be referred to as CB-1 and CB-2. Further gel filtration did not indicate any additional impurities in fraction CB-1. Thus, the actual yield of peptide CB-1 recovered from peak 1 is 0.75 μ mol. Re-chromatography of peak 2 indicated an insignificant contamination with peptide CB-1 which does not affect the calculated yield of this material. Only an insignificant amount of peptide CB-2 was detected in the two minor peaks. Thus, the quantities of peptides CB-1 and CB-2 recovered from peaks 1 and 2 are essentially stoichiometric. Such a result would be expected from the cyanogen bromide reaction.

A total of 0.89 μ mol of the cyanogen bromide digest was recovered out of the 1.18 μ mol loaded on to the column. This includes 0.12 μ mol of uncleaved fraction DU1A as well as the impurities in the two minor peaks. Thus, about 0.70–0.75 μ mol of cleaved peptides would be expected, in excellent agreement with the observed yield of peaks 1 and 2.

Based on gel electrophoresis in sodium dodecyl sulphate, peak 0 material has the same molecular weight as fraction DU1A. Amino acid analyses showed that it had the same composition as fraction DU1A, with the methionine residues replaced by homoserine. This material probably represents protein which was modified, but not cleaved.

When peak 1 and peak 2 were subjected to gel electrophoresis in sodium dodecyl sulphate, each gave a single component with only trace amounts of impurities (Figs. 2c and 2d). These components are identical with the major bands observed in the unfractionated digest (Fig. 2b). After a second gel

filtration, no impurities were detected on gel electrophoresis in sodium dodecyl sulphate. Molecular weights of 12 600 and 9000 were determined for peptides CB-1 and CB-2 from such experiments.

Based on amino acid analyses, peptide CB-1 contains 110 amino acids and has mol. wt. 12 654. Peptide CB-2 contains 81 residues and has mol. wt. 9326 (Table 4). These molecular weights agree closely with those obtained from gel electrophoresis in sodium dodecyl sulphate, and the sum of these molecular weights is equal to the molecular weight of fraction DU1A. Except for the single-residue differences in the content of lysine, glycine, leucine and homoserine, the sum of the compositions of peptides CB-1 and CB-2 accounts for the composition of fraction DU1A. Both tryptophan residues are found in peptide CB-2 and only one residue of tyrosine and five of phenylalanine are found in peptide CB-1. This variation in the aromatic amino acid content accounts for the differences in the extinctions of the two peptides.

A one-step Edman degradation indicated that serine was the only *N*-terminal amino acid for peptide CB-1 and only aspartic acid was found for peptide CB-2. These observations agree with the *N*-terminal residues detected in the unfractionated digest.

The positions of peptides CB-1 and CB-2 in the sequence of fraction DU1A may be inferred from the specificity of the cyanogen bromide reaction. Cyanogen bromide cleaves peptide bonds involving the carboxyl group of methionine and releases peptides with *C*-terminal homoserine. Two components containing homoserine were isolated from the Sephadex G-50 fractionation of the cyanogen bromide digest of fraction DU1A. One component was peptide CB-2 and the other was found in the low-molecular-weight fractions. The latter component was shown to be *N*-acetylhomoserine by t.l.c. in two solvent systems, butanol–pyridine–acetic acid–water (15:10:3:12, by vol.) and ethanol–acetic acid–water (4:1:1, by vol.). The homoserine was also identified by amino acid analysis after acid hydroly-

Table 4. *Amino acid compositions and molecular weights of peptides CB-1 and CB-2 and subunit DU1A*

Experimental details are given in the text.

Amino acid	Content (residues/mol)			
	Peptide CB-1	Peptide CB-2	Peptides CB1+CB2	Subunit DU1A
Lysine	10	2	12	11
Histidine	6	4	10	10
Arginine	10	6	16	16
Aspartic acid	9	5	14	14
Threonine	4	3	7	7
Serine	9	10	19	19
Glutamic acid	14	6	20	20
Proline	8	10	18	18
Glycine	7	3	10	9
Alanine	6	4	10	10
Valine	8	2	10	10
Methionine	—	—	—	2
Isoleucine	6	4	10	10
Leucine	7	9	16	17
Tyrosine	1	1	2	2
Phenylalanine	5	9	14	14
Tryptophan	0	2	2	2
Homoserine	0	1	1	—
Total no. of residues	110	81	191	191
Molecular weights				
From amino acid composition	12 654	9326	21 980	22 069
From gel electrophoresis in sodium dodecyl sulphate	12 600	9000	21 600	22 500

Table 5. *Sequence studies on peptide CB-1*

Experimental details are given in the text. Abbreviations for 'Identification' are as follows: V, t.l.c. in solvent V, (Jeppson & Sjoquist, 1967); D and E, t.l.c. in solvents D and E (Edman & Sjoquist, 1956); G.C., gas chromatography (Pisano & Bronzert, 1969); S, Sakaguchi reaction (Jepson & Smith, 1953); P, Pauly reaction (Sanger & Tuppy, 1951).

Step no.	Residue	Yield (%)	Identification
1	Ser	100	V, E, G.C.
2	Leu	90	V, D, G.C.
3	Thr	—	V, E
4	Lys	—	V, E
5	Asp	—	V, E, G.C.
6	Phe	93	V, D, G.C.
7	Asp	—	V, E
8	Glu	—	V, E, G.C.
9	Val	69	V, D, G.C.
10	Asn	—	V, D
11	Ile	67	V, D, G.C.
12	Asp	—	V, E
13	Val	49	V, D, G.C.
14	Ser	—	V, E, G.C.
15	His	49	P, E, G.C.
16	Phe	46	V, D, G.C.

Table 6. *Sequence studies on peptide CB-2*

Experimental details are given in the text. Abbreviations for 'Identification' are defined in Table 5.

Step no.	Residue	Yield (%)	Identification
1	Asp	32	V, E, G.C.
2	Ile	100	V, D, G.C.
3	Ala	94	V, D, G.C.
4	Ile	105	V, D, G.C.
5	Ser	40	V, E, G.C.
6	His	—	P, E
7	Pro	80	V, E, G.C.
8	Trp	—	V
9	Ile	96	V, D, G.C.
10	Arg	—	S
11	Pro	30	V, G.C.
12	Ser	33	V, E, G.C.
13	Phe	78	V, D, G.C.
14	Phe	82	V, D, G.C.
15	Ser	—	V, G.C.
16	Phe	80	V, D, G.C.
17	His	—	P, E

sis. Peptide CB-2 must therefore be derived from the *N*-terminal region of fraction DU1A and since peptide CB-1 contains no homoserine, it must be

Fig. 4. *N*-Terminal sequences of (a) subunit DU1A and (b) peptide CB-1. The sequence of subunit DU1A was derived from that of peptide CB-2.

derived from the *C*-terminal region. These observations also indicate that *N*-acetylmethionine occupies the *N*-terminal position. Carboxypeptidase B digestion of peptide CB-1 released lysine. Since lysine was the *C*-terminal residue in fraction DU1A, this confirms that peptide CB-1 is derived from the *C*-terminal region of fraction DU1A.

Thus, peptides CB-1 and CB-2 each contain a single molecular-weight species, have only one *N*-terminal residue, and the sum of their amino acid compositions show good correlation with that of fraction DU1A. These observations indicate that the peptides are sufficiently pure for sequence determination. Sequential degradations of peptides CB-1 and CB-2 were therefore performed with the Beckman Sequencer and the results are summarized in Tables 5 and 6 respectively. A problem encountered with the sequencer was the solubility of the peptides in the organic solvents used in the extraction steps. Because of this difficulty, the peptide gradually moved up the wall of the cup and eventually into the groove, thus making some of the material inaccessible to the reagents. This can be seen from the results obtained with peptide CB-1 (Table 6) where the yields of phenylthiohydantoin derivatives occasionally decrease rapidly.

Some overlapping was also encountered in the later steps. This could perhaps be due to some of the peptide moving up the wall of the cup and then being redissolved after missing one or more degradation cycles.

Because of the decreased yields of phenylthiohydantoin derivatives and the overlapping problem, unequivocal identification of the derivatives was possible only up to residue 16 in peptide CB-1 and residue 17 in peptide CB-2, although other residues were also identified. Each peptide yielded a unique sequence. No evidence of microheterogeneity was found, indicating that these peptides were derived from a single subunit or from subunits having identical sequences in positions 1–18 and 82–98. The sequences are presented in Fig. 4.

DISCUSSION

Our results suggest that α -crystallin represented by peak 2 from a column of Bio-Gel A-5m is composed of a minimum of four distinctly different

subunits. Two of the subunits, represented by DU1A and DU1B, contain no thiol groups. The third subunit, represented by DU2B and DU3, which have almost identical amino acid compositions, contains one thiol group/molecule. Subunit DU4, although obviously impure, must contain an additional component with two thiol groups/molecule. In each case, the amino acid composition of the subunit cannot be derived from any combination of the other fractions. Subunit DU1A was obtained with almost no impurities. Subunits DU1B, DU2B and DU3 contain minor impurities and DU4 has significant contamination.

After correction for recoveries, it can be concluded that approx. 30% of the subunits do not contain thiol groups (DU1A and DU1B), about 60% contain one thiol group/molecule (DU2B and DU3), and after correction for impurities, 5% contain two thiol groups/molecule (DU4). The average cysteine content for this distribution of subunits is 0.8 thiol group/molecule, in good agreement with the values obtained from the amino acid composition and the specific radioactivity of carboxymethylated α -crystallin respectively.

Based on the results of Schoenmakers *et al.* (1969b) and Schoenmakers, Gerding, & Bloemendal (1969a), α -crystallin is composed of two closely related acidic polypeptides as well as two basic polypeptides. Although the two acidic polypeptides were separated and shown to have identical amino acid compositions (Schoenmakers *et al.* 1969a) the evidence for the presence of two basic polypeptides is based only on isoelectric focusing and gel electrophoresis. The amino acid composition of subunit DU1A is similar to that of the unfractionated basic polypeptides found by Schoenmakers *et al.* (1969b) with a major difference only in the serine content. This observation suggests that subunit DU1A may contain two components, but no evidence to support such a concept has been observed. Subunit DU1B must therefore represent a basic subunit not previously isolated. The amino acid composition of the acidic polypeptides is similar to those of subunits DU2B and DU3 (Schoenmakers *et al.* 1969a). Thus, α -subunits DU2B and DU3 probably represent the two acidic polypeptides of α -crystallin. Schoenmakers *et al.* (1969b) reported that the acidic polypeptides contain 1.6–2.0 mol of thiol group/

23 000 g, whereas our results indicate 1 mol of thiol group/mol. Palmer & Papaconstantinou (1969) described the isolation of the acidic polypeptides by chromatography on DEAE-cellulose. The amino acid compositions found by them are in good agreement with those by Schoenmakers *et al.* (1969a) and the present work. However, their results indicate one thiol group/molecule, in agreement with our observations. The work of Palmer & Papaconstantinou (1969) also suggests that one of the acidic subunits is not synthesized directly but arises as an age-dependent modification of the synthesized acidic polypeptide.

Subunit fractions DU1B and DU4 have not previously been described. Thus, although this work confirms the presence of the so-called acidic and basic polypeptides, it would appear that there are at least two other polypeptides present in α -crystallin.

Hoenders, De Groot, Gerding & Bloemendal (1969) have studied the molecular weight of the subunits of α -crystallin with the Archibald approach-to-equilibrium method. They reported that after a 3-day incubation in the presence of sodium dodecyl sulphate, the subunit molecular weight decreased from about 20 000 to about 12 000. Recently they repeated these studies and reported a value of 11 500 for the molecular weight (De Groot, Hoenders, Gerding & Bloemendal, 1970). They also claimed that α -crystallin contained one acetyl group/10 900 g although earlier reports (Mok & Waley, 1968; Hoenders, Schoenmakers, Gerding, Tesser & Bloemendal, 1968) had indicated there was one acetyl group/21 000–24 000 g.

We have reported (Spector *et al.* 1971b) that after incubation in 2% sodium dodecyl sulphate at 37°C for 3 days, molecular weights of 19 500 and 22 500 were still found for the subunits of α -crystallin. No effect was observed after 24 h incubations in 1 M-hydroxylamine or exposure to pH 3 or 11. Long-term sedimentation-equilibrium studies in 5.5 M-guanidinium chloride or 7 M-urea also have failed to detect such low-molecular-weight components. The results described in the present paper support the observations reported in the preceding paper. Determination of the molecular weight of subunit DU1A by gel electrophoresis in sodium dodecyl sulphate indicates that it belongs to the species of mol.wt. 22 500. Calculation of the molecular weight from the amino acid composition, based on two methionine residues per molecule gives a value of 22 069. After cleavage with cyanogen bromide, only two major peptides, in approximately equal amounts, were detected on the basis of gel electrophoresis and end-group analyses. The sum of the molecular weights of the cyanogen bromide peptides based on gel electrophoresis in sodium dodecyl sulphate and amino acid analyses is in good agree-

ment with the determined value for the unhydrolysed subunit.

It can be argued that subunit DU1A is atypical and that the other subunits actually have a molecular weight of about 11 000. However, subunit DU1A is a major constituent of α -crystallin, represents more than 20% of the total material and undoubtedly accounts for much more than half of the so-called basic peptides. Leon, De Groot & Bloemendal, (1970) claim that they have obtained the same low molecular weight with the basic peptides as previously observed with the total α -crystallin preparation (De Groot *et al.* 1970). The DU1A species should have made a major contribution in such experiments. The major species of all other isolated subunits had mol.wt. 22 500 and 19 000 on sodium dodecyl sulphate gels. The reason for the difference between the results of the two laboratories is not apparent. In our experiments, small amounts of low-molecular-weight components have been observed in the unfractionated carboxymethylated α -crystallin. Occasionally these components appear to increase in quantity in some of the isolated subunits, but they always represent a minor fraction of the total material. Further investigation will be required to resolve this problem.

The amino acid compositions indicate that the subunits of α -crystallin are related to each other. Corran & Waley (1969) determined the *N*-terminal sequences of unfractionated acidic and basic polypeptides isolated by the procedure of Schoenmakers & Bloemendal (1968). Almost identical results were obtained for the sequences. The *N*-terminal sequence of the A chain was found to be:

N-Acetyl-Met-Asp-Ile-Ala-Ile-Gln-His-Pro-Trp-Phe-Lys.

This may be compared with the corresponding sequence of subunit DU1A:

N-Acetyl-Met-Asp-Ile-Ala-Ile-Ser-His-Pro-Trp-Ile-Arg.

Eight of the eleven positions have identical residues, and conservative substitutions are found in positions 6, 10 and 11. Corran & Waley's (1969) sequence of the *N*-terminal region of the basic chains would be identical with that of subunit DU1A if they had found a serine residue in position 6 between the isoleucine and histidine residues. It is likely that they missed this serine residue since their sequence studies were performed on about 0.1 μ mol of peptide and most amino acid analyses were done qualitatively by paper chromatography.

There is a preponderance of hydrophobic amino acids in the *N*-terminal region of subunit DU1A. Eleven of the first 18 residues are hydrophobic.

Only four of the first 18 residues are charged and besides this, the α -amino group is acetylated. Peptide CB-2 contains only 23 amino acids capable of carrying a charge, whereas 51 are found in peptide CB-1. Thus, it appears that the charged amino acids are concentrated in the C-terminal region of subunit DU1A.

It is not possible to disaggregate α -crystallin completely in acid or alkali or by modification of the ϵ -amino group of lysine with succinic anhydride (Spector & Katz, 1965). Hence it seems likely that hydrophobic interactions play an important role in the stabilization of the aggregate. Perhaps this is reflected in the apparent separation of the hydrophobic and electrostatic groups in the DU1A polypeptide chain.

It has been demonstrated (Spector, Li, Augusteyn & Meretsky, 1971a) that subunit DU1A is capable of forming a macromolecule in the size range of α -crystallin. Hence, there is no requirement for the presence of acidic chains to produce an aggregated macromolecule. There is also no apparent requirement for thiol groups. The similarities in the amino acid compositions, as well as the N-terminal sequences, suggest that these subunits may have evolved from a common ancestral form. The diversity of polypeptides may therefore be a result of evolutionary change and not a reflexion of physiological requirement.

This work was supported by The John A. Hartford Foundation, Inc. and The National Eye Institute, National Institutes of Health, U.S. Department of Health, Education and Welfare. We are pleased to acknowledge the competent technical assistance of Mrs Faith Hughes and Miss Halina Szejnwald. We also thank the Fight For Sight Inc. for their support.

REFERENCES

- Björk, I. (1964). *Expl Eye Res.* **3**, 239.
 Corran, P. H., & Waley, S. G. (1969). *Biochem. J.* **115**, 789.
 Crestfield, A. M., Moore, S. & Stein, W. H. (1963). *J. biol. Chem.* **238**, 622.
 De Groot, K., Hoenders, H. J., Gerding, J. J. T. & Bloemendal, H. (1970). *Biochim. biophys. Acta*, **207**, 202.
 Edelhoch, H. (1967). *Biochemistry, Easton*, **6**, 1948.
 Edman, P. & Begg, G. (1967). *Eur. J. Biochem.* **1**, 80.
 Edman, P. & Sjoquist, J. (1956). *Acta chem. scand.* **10**, 1507.
 Gray, W. R. & Hartley, B. S. (1963). *Biochem. J.* **89**, 379.
 Hoenders, H. J. & Bloemendal, H. (1967). *Biochim. biophys. Acta*, **147**, 183.
 Hoenders, H. J., De Groot, K., Gerding, J. J. T. & Bloemendal, H. (1969). *Biochim. biophys. Acta*, **188**, 162.
 Hoenders, H. J., Schoenmakers, J. G. G., Gerding, J. J. T., Tesser, G. I. & Bloemendal, H. (1968). *Expl Eye Res.* **7**, 291.
 Jeppson, J. O. & Sjoquist, J. (1967). *Analyt. Biochem.* **18**, 264.
 Jeppson, B. J. & Smith, I. (1953). *Nature, Lond.*, **172**, 1100.
 Leon, A. E., De Groot, K. & Bloemendal, H. (1970). *Expl Eye Res.* **10**, 75.
 Mok, C. C. & Waley, S. G. (1968). *Expl Eye Res.* **7**, 148.
 Morse, D. & Horecker, B. L. (1966). *Analyt. Biochem.* **14**, 429.
 Palmer, W. G. & Papaconstantinou, J. (1968). *Biochemistry, Easton*, **7**, 243.
 Palmer, W. G. & Papaconstantinou, J. (1969). *Proc. natn. Acad. Sci. U.S.A.* **64**, 404.
 Pisano, J. & Bronzert, T. J. (1969). *J. biol. Chem.* **244**, 5597.
 Sanger, F. & Tuppy, H. (1951). *Biochem. J.* **49**, 463.
 Schoenmakers, J. G. G. & Bloemendal, H. (1968). *Biochem. biophys. Res. Commun.* **31**, 257.
 Schoenmakers, J. G. G., Gerding, J. J. T. & Bloemendal, H. (1969a). *Eur. J. Biochem.* **11**, 472.
 Schoenmakers, J. G. G., Matze, R., Van Poppel, M. & Bloemendal, H. (1969b). *Int. J. Protein Res.* **1**, 19.
 Shapiro, A. L., Vinuela, E. & Maizel, J. V., jun. (1967). *Biochem. biophys. Res. Commun.* **28**, 815.
 Spande, T. F. & Witkop, B. (1967). In *Methods in Enzymology*, vol. II, p. 498. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
 Spector, A. (1965). *Investve Ophthal.* **4**, 579.
 Spector, A. (1970). *Proc. 8th int. Cong. Biochem., Switzerland* (in the Press).
 Spector, A. & Katz, E. (1965). *J. biol. Chem.* **240**, 1979.
 Spector, A., Li, L.-K., Augusteyn, R. C. & Meretsky, D. (1971a). *Am. J. Ophthal.* **71**, 386.
 Spector, A., Li, L.-K., Augusteyn, R. C., Schneider, A. & Freund, T. (1971b). *Biochem. J.* **124**, 337.
 Spector, A., Wandel, T. & Li, L.-K. (1968). *Investve Ophthal.* **7**, 179.
 Weber, K. & Osborn, M. (1969). *J. biol. Chem.* **244**, 4406.