

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/142125>

Please be advised that this information was generated on 2021-10-26 and may be subject to change.

## The Amino-Acid Sequence of the $\alpha A_2$ Chain of Bovine $\alpha$ -Crystallin

Frans J. VAN DER OUDERAA, Wilfried W. DE JONG, and Hans BLOEMENDAL

Laboratorium voor Biochemie, Universiteit van Nijmegen

(Received May 23, 1973)

The  $\alpha A_2$  chain of bovine  $\alpha$ -crystallin was fragmented by means of cyanogen bromide treatment and by tryptic, chymotryptic and thermolytic digestions. Twenty tryptic peptides were obtained from the S-aminoethylated  $\alpha A_2$  chain, accounting together for the complete sequence. The direct Edman degradation and the dansyl-Edman technique were used to determine the sequences of the tryptic peptides. The order of the tryptic peptides was deduced from overlapping peptides obtained by cyanogen bromide treatment, tryptic digestion of the maleylated chain and chymotryptic and thermolytic digestion of the S-aminoethylated chain. The sequence of the  $\alpha A_2$  chain comprises 173 residues and corresponds to a molecular weight of 19832.

$\alpha$ -Crystallin is one of the major water-soluble proteins of the vertebrate eye lens. It is considered as a structural protein, and is apparently devoid of any biological activity.  $\alpha$ -Crystallin is isolated as an aggregate with an average molecular weight of 800000 [1] and is composed of two different types of polypeptide chains,  $\alpha A$  and  $\alpha B$ , with molecular weights of approximately 20000 [2,3]. Each of these chains occurs in two electrophoretically distinguishable forms:  $\alpha A_1$  and  $\alpha A_2$ , and  $\alpha B_1$  and  $\alpha B_2$ , respectively. The  $\alpha A$  chains have identical amino acid compositions [2], and  $\alpha A_1$  is supposed to be derived from  $\alpha A_2$  by deamidation *in vivo* [4,5].

Partial sequences of the bovine  $\alpha A$  chain have been published previously, *viz.* the N-terminal 13 residues [6,7], 28 residues around the only cysteine [8] and the C-terminal cyanogen bromide fragment [9]. The N-terminus is blocked by acetylation [10].

Considerable interest has recently been focused on  $\alpha$ -crystallin since a messenger-RNA fraction has been isolated from calf lenses which codes exclusively for the  $\alpha A_2$  chain as shown by translation *in vivo* and *in vitro* in heterologous systems [11—13].

In this paper the complete sequence of the 173 residues in the  $\alpha A_2$  chain is presented, as deduced from the sequences of 20 tryptic peptides. Overlaps of these peptides after maleylation and tryptic digestion of the whole polypeptide and by digestion with other proteolytic enzymes established the proposed structure.

**Abbreviations.** Dansyl, 5-dimethylaminonaphthalene-1-sulphonyl; CNBr, cyanogenbromide; Aet-, S-aminoethyl-

**Enzymes.** Aminopeptidase M (EC 3.4.1.2); carboxypeptidase A (EC 3.4.2.1); carboxypeptidase C (EC 3.4.2.—); trypsin (EC 3.4.4.4); chymotrypsin (EC 3.4.4.5); thermolysin (EC 3.4.24.4); pepsin (EC 3.4.4.1); pronase (EC 3.4.4.—).

### EXPERIMENTAL PROCEDURE

#### MATERIALS

Calf eyes were freshly obtained from the slaughter house and the lenses removed and processed as soon as possible.

Silica-gel thin-layer chromatography plates (G1500 LS 254) and polyamide thin-layer sheets (F1700) were purchased from Schleicher and Schüll, Aminex A-5 resin from Bio-Rad Laboratories and Sephadex from Pharmacia. Whatman 3 MM paper was used for electrophoresis and chromatography. 3% Dextrin GC 300 on Supelcoport, 100—120 mesh, was obtained from Supelco (Bellefonte, Pa.), 3% OV-17 on Chromosorb W-HP, 100—120 mesh, from Chrompack (Vlissingen, The Netherlands), *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide from Machery and Nagel and sodium 4-sulfophenylisothiocyanate monohydrate (Sequenal grade) from Pierce.

Pyridine and dimethylallylamine, used for Edman degradation were distilled twice after refluxing with KOH and phthalic anhydride, respectively, over a spinning-bend distilling column. Phenylisothiocyanate was distilled under reduced pressure as described by Edman and Begg [14]. Ethyl acetate, butyl chloride and benzene were distilled over CaH<sub>2</sub> before use. All other reagents were of analytical grade or better.

Trypsin (treated with L-1-tosylamido-2-phenylethyl chloromethyl ketone),  $\alpha$ -chymotrypsin, pepsin and carboxypeptidase A (diisopropylphosphorofluoridate-treated) were from Worthington, thermolysin from Calbiochem, carboxypeptidase C and aminopeptidase M from Röhm and Haas and pronase P from Serva.

## METHODS

### Isolation of $\alpha A_2$

Crude  $\alpha$ -crystallin was prepared by isoionic precipitation of calf-lens cortex extract in the presence of  $ZnSO_4$  and ethanol [15]. This material was purified by zonal centrifugation and appeared to be homogeneous as judged by alkaline polyacrylamide-gel electrophoresis in the presence of 7 M urea, immunoelectrophoresis and amino-acid analysis [1]. The  $\alpha A_2$  chain was isolated from  $\alpha$ -crystallin by ion-exchange chromatography on SE-Sephadex C-50 at pH 5.5 in 7 M urea [2] and showed a single band on electrophoresis. All procedures were carried out at 4 °C.

### Reduction and Aminoethylation

Up to 100 mg  $\alpha A_2$  or cyanogen-bromide-treated  $\alpha A_2$  was reduced in 7.5 ml 0.2 M Tris-HCl buffer pH 8.6, containing 15 mg of the sodium salt of EDTA, 4.3 g guanidinium chloride and 0.1 ml 2-mercaptoethanol. After one hour, 0.2 ml ethyleneimine was added and allowed to react for 2 h at room temperature [16]. The mixture was desalted on a Sephadex G-25 column, equilibrated with 0.1 M ammonia and lyophilized.

### Preparation of Cyanogen-Bromide Fragments

Cyanogen bromide cleavage [19] was performed in 70% formic acid at a protein concentration of 10 mg/ml, using an equal weight of CNBr. After 18 h at room temperature in the dark the solution was diluted with 10 volumes of water and concentrated by rotary evaporation. The residue was again diluted and lyophilized.

The cleaved protein was reduced and aminoethylated and subsequently desalted and lyophilized. This material was dissolved in 1 ml 1% ammonia and applied to a column of Sephadex G-50 fine (120 × 1.5 cm), equilibrated with 0.1 M ammonia at room temperature and eluted at a flow rate of 7 ml/h. The effluent was monitored at 280 and 230 nm.

### Blocking of $\epsilon$ -NH<sub>2</sub> Groups of Lysine

Maleylation of the  $\epsilon$ -NH<sub>2</sub> groups of lysine was carried out as described by Butler *et al.* [17], using 0.2 M sodium borate buffer pH 9.0 and a 60-fold molar excess of maleic anhydride over total protein amino groups; the reaction temperature was kept at 4 °C. Excess reagents were removed by gel filtration over Sephadex G-25 in 0.1 M ammonia.

Deblocking of maleylated peptides was performed as described by Grosclaude [18].

### Enzymic Digestions

Tryptic digestion was carried out in 0.1 M ammonium bicarbonate buffer pH 8.9 at a protein concen-

tration of 10 mg/ml. Trypsin (1%, w/w) was added at zero time and again after one hour. After 3 h at 37 °C the pH was lowered with 1 N HCl to 3.1 and the digest was heated for 5 min at 100 °C. The insoluble core was removed by centrifugation and washed twice with pH 3.1 buffer. The combined supernates were lyophilized.

Chymotryptic and thermolytic digestions were performed at 37 °C in 0.1 M ammonium bicarbonate buffer pH 8.9 and 8.6, respectively. The enzyme/substrate ratio was 0.5% (w/w) in the case of  $\alpha A_2$ , and 0.5 mg/ $\mu$ g in the case of peptides, using digestion times of 1 h and 15 h, respectively.

Peptic digestion was carried out in 0.01 M HCl for 15 h at room temperature, using 0.3 mg pepsin per  $\mu$ mol peptide. Complete enzymic hydrolysis of peptides was obtained by adding 0.04 mg pronase and 0.04 mg aminopeptidase M in 200  $\mu$ l 0.05 M sodium borate buffer pH 8.0, to about 50 nmol peptide. After 16 h at 30 °C the mixture was diluted with 0.8 ml sodium citrate buffer pH 2.2, and used directly for amino acid analysis. Carboxypeptidase A digestion was carried out as described earlier [2] and carboxypeptidase C digestion was done according to Tschesche and Kupfer [20].

### Peptide Separation

Analytical peptide mapping of enzymic digests was carried out by high-voltage paper electrophoresis in pyridine—acetic acid—water buffer (25:1:225, by vol.; pH 6.5) in a varsol-cooled apparatus at 45 V/cm, followed in the second dimension by descending paper chromatography in *n*-butanol—acetic acid—water—pyridine (15:3:12:10, by vol.) [21]. Staining was performed with 0.02% ninhydrin in acetone, containing 1% pyridine and 1% acetic acid, or with selective staining procedures [22].

For preparative purposes, 1.5–2.0  $\mu$ mol of the protein digest was applied as a band of 20 cm on Whatman 3 MM paper and electrophoresed at pH 6.5. Poorly resolved zones were sometimes submitted to a second electrophoretic run at pH 3.8 (pyridine—acetic acid—water, 3:20:377, by vol.). The electrophoretically separated zones were further purified by descending chromatography in the above mentioned solvent. Peptides were located by staining guide strips with 0.5% ninhydrin in acetone. The unstained material was eluted with 10% acetic acid (in case of basic and neutral peptides) or 0.1 M ammonia (for acidic peptides) and taken to dryness.

Alternatively the soluble tryptic peptides of  $\alpha A_2$  were fractionated by ion-exchange chromatography. A column (15 × 0.4 cm) of Aminex A-5 was equilibrated at 50 °C with 0.2 M pyridine acetate pH 3.1, and loaded with 2.0  $\mu$ mol digest. Elution of peptides was accomplished with a linear gradient from 0.2 M pyridine acetate pH 3.1, to 2.0 M pyridine acetate

pH 5.0, using 125 ml of each buffer [23]. To detect the peptide fractions small aliquots of every second tube were spotted on Whatman 3 MM paper. After descending chromatography the paper was stained with 0.5% ninhydrin in acetone or the Reindel-Hoppe reagent [24]. Further purification of the peptide fractions was obtained by descending paper chromatography. Tryptic peptides, insoluble at pH 3.1, were redissolved in 1 ml 1% ammonia and applied to a column (120×1.5 cm) of Sephadex G-50 fine equilibrated at room temperature with 0.1 M ammonia. The flow rate was 7 ml/h and the effluent was monitored at 230 nm.

Chymotryptic, peptic and thermolytic peptides were in all cases isolated by high-voltage paper electrophoresis and descending chromatography.

#### *Amino-Acid Analysis*

Peptide samples (10–50 nmol) were hydrolysed in 0.5 ml 6 N HCl for 22 h at 110 °C in evacuated tubes. In case of peptides containing tyrosine, 0.1% (w/v) phenol was added to the 6 N HCl. No corrections were made for hydrolytic losses of threonine and serine. The composition of  $\alpha A_2$  and its CNBr fragments was obtained from duplicate analysis after 24, 48 and 72-h hydrolyses. Values for serine and threonine were extrapolated back to zero time and the values for valine and isoleucine were taken from the 72-h hydrolysate.

Analysis were performed with a Beckman Multi-chrom amino-acid analyser equipped with a 16-mm cuvette, using the single column method as described in the manufacturer's manual. The tryptophan content of  $\alpha A_2$  was determined spectrophotometrically on unhydrolysed samples according to Edelhoeh [25]. Ehrlich's reagent was used to demonstrate the presence of tryptophan in peptides.

#### *N-Terminal Sequence Analysis*

The three stage Edman degradation technique was used essentially as described by Niall [26]. For Edman degradation of some of the larger lysyl peptides, 4-sulfophenylisothiocyanate was used at the first step [27], performing the coupling of peptides in a buffer 0.4 M in dimethylallylamine and 0.4 M in dimethylamino-1-propyne, adjusted to pH 9.5 with trifluoroacetic acid. After two hours incubation at 50 °C the 4-sulfophenylthiocarbonyl peptide was extracted four times with butyl chloride, lyophilized and subsequently subjected to the first step of the normal phenylisothiocyanate coupling procedure. The dansyl-Edman procedure was carried out as described by Gray and Smith [28].

#### *Identification of N-Terminal Residues*

The amino acid phenylthiohydantoin derivatives were identified both by thin-layer chromatography

and by gas chromatography. The solvents used for silica-gel thin-layer chromatography were: chloroform–methanol (9:1, by vol.) [29] and heptane–propionic acid–dichloroethane (58:17:25, by vol.) [30].

Gas chromatography was performed at three isothermal conditions in order to come to an unambiguous determination of the phenylthiohydantoin derivatives [31]. The identity of the apolar and hydroxylic amino acid derivatives was established at 190 °C on a column (6 ft×2 mm), containing 3% Dexsil 300 GC. These derivatives of phenylalanine, asparagine, glutamine and tyrosine, as well as the silylated derivatives of glutamic and aspartic acid, were chromatographed on a column (4 ft×4 mm) of 3% OV-17, at oven temperatures of 275 °C and 220 °C, respectively. Silylation was carried out in equal volumes of *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide and ethyl acetate for 30 min at room temperature. The phenylthiohydantoin derivatives of alanine, glycine, valine, proline, leucine, isoleucine, phenylalanine and methionine, were determined quantitatively. The histidine and arginine derivatives were identified by the Pauly and phenantrenequinone spot tests, respectively. Identification of dansyl amino acids was established by thin-layer chromatography on polyamide sheets (5×5 cm) [32].

#### *Amide Assignment*

Amide groups were determined both by identifying the phenylthiohydantoin derivatives of glutamate and aspartate after direct Edman degradation and by comparing electrophoretic mobilities of peptides at pH 6.5 [33]. In addition complete enzymic hydrolysis, followed by amino acid analysis was used in some cases.

#### *Peptide Nomenclature*

Peptides are numbered on the basis of their positions in the complete sequence, starting from the N-terminus. Tryptic peptides are preceded by T, chymotryptic by C, thermolytic by Th and peptic by P. CB refers to cyanogen bromide fragments and Tm to tryptic peptides obtained after maleylation. Peptides derived by further cleavage are indicated by a subscript to the parent peptide.

## RESULTS

### *Cyanogen-Bromide Peptides*

Cyanogen bromide treated  $\alpha A_2$  chain was amino-ethylated and thereafter fractionated as shown in Fig. 1. The  $\alpha A_2$  chain contains two methionyl residues; one is the N-terminal acetylated residue, while the other one occupies an internal position. The peaks marked I and II in Fig. 1 contain the two

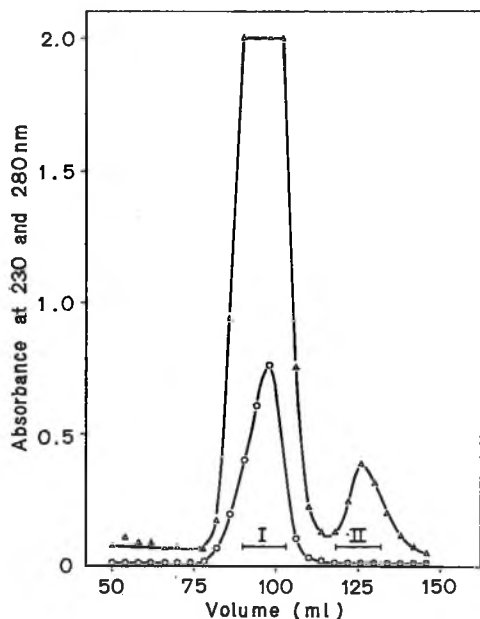


Fig. 1. Fractionation of CNBr fragments of S-aminoethylated  $\alpha A_2$ . 30 mg CNBr peptides were dissolved in 1 ml 1% ammonia and separated on a column of Sephadex G-50 fine (120  $\times$  1.5 cm). Elution at room temperature was performed with 0.1 M ammonia at a flow rate of 7 ml/h. The volume of each fraction was 4 ml. ( $\Delta$ ) absorbance at 230 nm; ( $\circ$ ) absorbance at 280 nm

CNBr fragments of the  $\alpha A_2$  chain, expected on the basis of one internal methionyl residue. The results of amino acid analysis and end-group determination of the two fragments and of  $\alpha A_2$  are summarized in Table 1. The N-terminal acetylated methionine, which is also split off by the action of CNBr, is not considered as a separate CNBr peptide and therefore not mentioned in the table.

It was concluded that fraction I contained the N-terminal fragment  $\alpha A_2$ CB1, on the basis of the following properties. First, its N-terminal residue was aspartic acid, which is the second residue of the  $\alpha A_2$  chain, next to the N-terminal methionine [10]. Second, the only cysteinyl residue, which precedes the internal methionine [8], is present in this fragment. Third, homoserine lactone derived from the internal methionyl residue is found in this fragment. Fraction II contains the C-terminal fragment  $\alpha A_2$ -CB2, of which the sequence has been published previously [9]. The apparent low absorbance at 280 nm of this fragment is consistent with the absence of tryptophanyl and tyrosinyl residues.

#### Tryptic Peptides of S-Aminoethylated $\alpha A_2$

Fingerprinting of the tryptic peptides, soluble at pH 3.1, of the aminoethylated  $\alpha A_2$  chain (Fig. 2) indicated that most peptides could easily be isolated

Table 1. Amino-acid compositions of the  $\alpha A_2$  chain and its cyanogen-bromide fragments

Data are given as numbers of residues per protein molecule as determined by amino acid analysis and sequence results

Amino acid	$\alpha A_2$		CB1		CB2	
	composition	sequence	composition	sequence	composition	sequence
Aspartic acid	16.1	16	14.8	15	1.1	1
Threonine	5.0	5	3.9	4	1.0	1
Serine	23.3	23	15.3	15	7.8	8
Glutamic acid	17.7	17	14.2	14	3.1	3
Proline	12.1	12	7.1	7	5.0	5
Glycine	10.5	10	7.4	7	3.1	3
Alanine	6.2	6	3.2	3	3.1	3
Valine	10.0	10	8.0	8	2.0	2
Cysteine	+	1	+	1	—	—
Methionine	1.8	2	—	—	—	—
Isoleucine	9.1	9	6.9	7	1.9	2
Leucine	14.3	14	13.0	13	1.0	1
Tyrosine	6.1	6	5.9	6	—	—
Phenylalanine	13.9	14	12.8	13	1.0	1
Lysine	7.1	7	5.1	5	1.9	2
Histidine	7.0	7	5.8	6	1.0	1
Arginine	13.1	13	11.1	11	2.0	2
Homoserine lactone	—	—	+	1	—	—
Tryptophan	1.3	1	+	1	—	—
Total		173		137		35
N-Terminal		Blocked		Asp		Leu

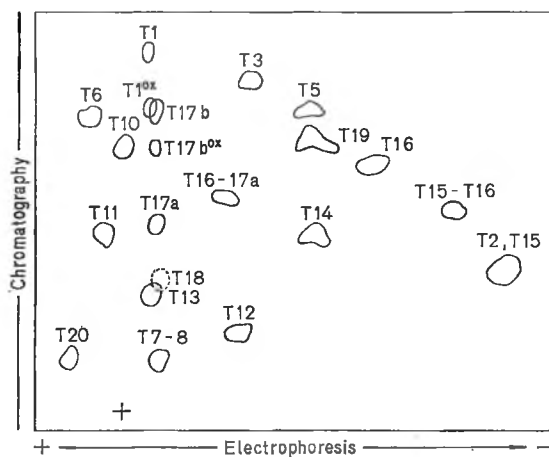


Fig. 2. Peptide map of a tryptic hydrolysate of aminoethyl- $\alpha A_2$ . High-voltage electrophoresis (pH 6.5) was carried out at 50 V/cm for 90 min and followed by descending chromatography for 18 h in *n*-butanol–acetic acid–water–pyridine (15:3:12:10, by vol.). Peptides were located with 0.02% ninhydrin in buffered acetone

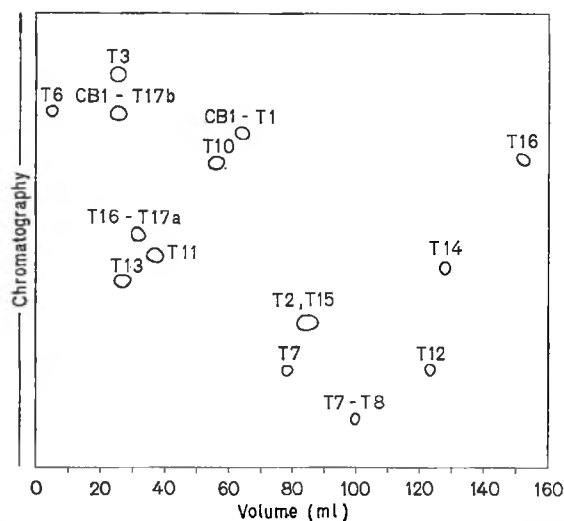


Fig. 3. Fractionation of tryptic peptides of aminoethyl-CB1. The soluble tryptic peptides from 30 mg aminoethyl-CB1 were chromatographed on a column (15  $\times$  0.4 cm) of Aminex A-5 equilibrated at 50  $^{\circ}$ C with 0.2 M pyridine acetate pH 3.1. Elution was performed at a flow rate of 12 ml/h using a linear gradient from 0.2 M pyridine acetate pH 3.1 to 2.0 M pyridine acetate pH 5.0. The fraction volume was 1.0 ml. 10% of every second fraction was concentrated and applied on a sheet of Whatman 3 MM paper for descending chromatography. The paper was stained with 0.5% ninhydrin in acetone

by high-voltage electrophoresis at pH 6.5, followed by descending chromatography. Likewise, a good separation of the soluble tryptic peptides could be obtained by column chromatography on Aminex A-5 (Fig. 3). The "core" material, insoluble at pH 3.1

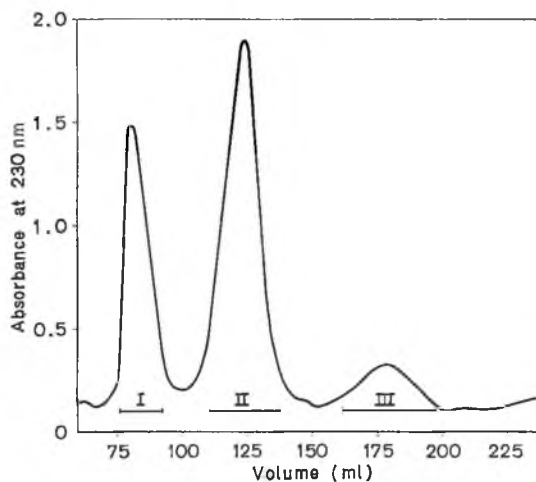


Fig. 4. Fractionation of pH-3.1-insoluble tryptic peptides. The material insoluble at pH 3.1, obtained after tryptic digestion of 1.5  $\mu$ mol  $\alpha A_2$ , was dissolved in 1% ammonia and separated on Sephadex G-50 fine under the conditions described in Fig. 1. The fractions I, II and III contain undigested  $\alpha A_2$ , peptide T4 and peptide T9, respectively

was fractionated by gel chromatography in 0.1 M ammonia on Sephadex G-50 fine (Fig. 4) and yielded two pure peptides, T4 and T9, in addition to the remaining undigested material. The isolation of T4 succeeded only when the core of S-aminoethylated  $\alpha A_2$  was used. In case of tryptic digests of native or carboxymethylated  $\alpha A_2$ , the core contained also peptide T17a–17b, which eluted from the Sephadex column together with T4.

The composition of all the tryptic peptides of S-aminoethylated  $\alpha A_2$  is given in Table 2, together with the isolation procedures and N-terminal determinations. Some of these peptides were further degraded by digestion with pepsin, thermolysin or chymotrypsin (Table 3) in order to facilitate their sequence determination. The sequences of all tryptic peptides were established as shown in Table 4.

The sequence of T1 could not be determined directly, since it is blocked by the N-terminal acetyl group, and was therefore deduced from the first 10 degradation steps on  $\alpha A_2$ CB1 and from the thermolytic peptides obtained from CB1T1.

T2 and T15 were characterized as free arginine, rather than Arg-Arg dipeptides, by dansylation of these peptides, followed by polyamide thin-layer chromatography without prior acid hydrolysis. The dansylated peptides moved in the position of dansyl-arginine. The sequence results of overlap peptides (to be discussed later) shows indeed that two free residues of arginine must be released by tryptic digestion of the  $\alpha A_2$  chain.

The sequence of the large and hydrophobic peptide T4 could be established conclusively up to residue 19 by the direct Edman method. Dansyl-

Table 2. *Amino-acid compositions of the*  
 Data are given as molar ratios. No corrections for hydrolytic destruction have been made. Purification procedures were (see filtration on Sephadex G-50 fine (S). Electrophoretic mobility

Amino acid	Peptide										
	T1	T2	T3 <sup>a</sup>	T4 <sup>a</sup>	T5	T6	T7-8	T9 <sup>b</sup>	T10	T11	T12
Aspartic acid	0.97			2.17		1.13	2.07	0.98	1.03	1.06	1.07
Threonine			1.01	0.78		0.83			0.90		
Serine			0.99	2.47	0.81	1.95	0.98		1.03		
Glutamic acid	0.95			3.12	1.05	1.11			1.05	3.09	1.12
Proline	1.12		2.05	2.00					1.13		
Glycine			1.02	2.09		1.01				1.10	
Alanine	0.94										
Valine						1.96		1.77	1.07	2.02	
Methionine	0.98										
Isoleucine	2.06			1.04		1.17		1.02		0.87	
Leucine			1.00	5.14	0.95	1.01		1.06	1.05		
Tyrosine			0.84	2.95							
Phenylalanine	1.00		1.00	5.16	1.06			2.12	0.91	0.86	
Lysine	1.00							1.00	1.05	1.03	1.06
Aminoethyl-cysteine											
Histidine	0.97								0.80	0.94	0.77
Arginine		1.00	0.93	0.92	1.13	0.90	0.95				1.04
Tryptophan <sup>c</sup>	+										
No. of residues	11	1	9	28	5	11	5	8	10	11	4
N-terminal residue	—	Arg	Thr	Leu	Gln	Thr	Ser	Phe	His	Val	His
Mobility	0	1.00	0.25	0	0.43	-0.21	0	0	-0.11	-0.16	0.23
Purification	E-C	E	E-C A-C	S	E-C	E-C	E-C A	S	E-C A	E-C A	E-C A

<sup>a</sup> Hydrolysed in the presence of 9 mg phenol per 100 ml 6 N HCl.

<sup>b</sup> Hydrolysed for 72 h.

<sup>c</sup> Determined with Ehrlich reagent.

Edman degradation of four chymotryptic peptides from T4 established the remaining sequence of this peptide, which was in part confirmed by some thermolytic peptides.

The N-terminal glutamyl residue of T5, as well as of T13, did easily cyclize to pyrrolidone carboxylic acid, which seriously interfered with the sequence determination.

The Arg-Asp bond between peptides T7 and T8 was cleaved very incompletely by trypsin, even after prolonged digestion. Consequently the combined peptide, T7-8, was isolated and sequenced. Peptides

T10, T11 and T17a were coupled with 4-sulfophenylisothiocyanate before starting the direct Edman degradation. This modification considerably improved the repetitive yields, since peptide losses due to extraction into the organic phase were diminished. The uncyclized form of T13 was isolated in insufficient amounts for direct Edman degradation. Therefore its sequence was derived from dansyl-Edman degradation of the peptic peptides T13P1 and T13P2 and sustained by the composition of T13C1 and T13C2. The sequence of the methionine containing peptide T17b was established by eight degradation

*tryptic peptides from aminoethylated  $\alpha A_2$  chain*

text for details): electrophoresis at pH 6.5 (E); paper chromatography (C); column chromatography on Aminex A-5 (A); gel was taken at pH 6.5, relative to arginine (= 1.00)

Amino acid	Peptide										
	T13*	T14	T15	T15-16*	T16*	T16-17a*	T17a	T17b	T18	T19	T20
Aspartic acid	1.90					2.11	2.00	0.99	1.00		
Threonine								0.79			
Serine	0.93					2.85	2.54	2.72	1.72	0.98	3.59
Glutamic acid	1.00	1.01				1.18	1.09		1.01		2.04
Proline						0.84	1.03	1.18	1.10	1.04	2.16
Glycine	1.02							1.96	2.00		
Alanine						1.15	1.03	1.10	1.08	0.97	1.09
Valine						1.07	1.11		0.99	0.99	
Methionine								0.75			
Isoleucine	1.03								1.05	1.01	
Leucine						1.78	2.22	2.05			
Tyrosine	0.79		0.93	0.76	0.73						
Phenylalanine		0.92						0.96			
Lysine								1.04			1.12
Aminoethyl-cysteine						+	+				
Histidine	0.97	0.90							0.99		
Arginine	1.14	1.17	1.00	2.07	1.24	1.12			1.07	1.01	
No. of residues	9	4	1	3	2	14	12	14	12	6	10
N-terminal residue	Gln	Glu	Arg	Arg	Tyr	Tyr	Leu	Ser	Ile	Ala	Glu
Mobility	0	0.45	1.00	0.85	0.62	0.20	0	0	0	0.46	-0.26
Purification	A-C	E-C A	E	E	E	E-C	E-C A-C	E-C	A	E-C	E-C

steps on T17b itself, combined with the results of the dansyl-Edman degradation of CB2T1, the N-terminal tryptic peptide of CB2.

Special attention was paid to the elucidation of the sequence of T20, which proved to be the C-terminal peptide. Eight residues could be sequenced by the direct Edman method. Digestion with carboxypeptidase A released a single residue of serine and treatment with carboxypeptidase C for 2.5 h released 1.9 residues of serine, 1.0 of proline and 0.4 of alanine. The sequence determination of the two thermolytic peptides T20Th1 and T20Th2 finally affirmed the total sequence of T20.

*Tryptic Peptides of Maleylated  $\alpha A_2$ CB1*

Aminoethylated  $\alpha A_2$ CB1 was maleylated and subsequently hydrolysed with trypsin. A small part of the digest was deblocked to liberate the  $\epsilon$ -amino groups and fingerprinted as shown in Fig. 5. Comparison with the fingerprint of aminoethylated  $\alpha A_2$  (Fig. 2) shows that all lysyl peptides of CB1, namely T1, T7-8, T9, T10 and T11, have selectively disappeared. T12 is the only arginyl peptide that has been lost. The apparent absence of T13 is presumably due to cyclization of its N-terminal glutamine.

The only new spot that has appeared, CB1Tm1, showed fluorescence under ultraviolet light, indicat-



Table 3. Amino-acid compositions of peptides obtained by thermolytic (*Th*), peptic (*P*), or chymotryptic (*C*) digestion of some tryptic peptides  
 CB2T1 is the N-terminal tryptic peptide of  $\alpha_2$ CB2. Data are given as molar ratios. These peptides were all purified by paper electrophoresis at pH 6.5, followed by descending chromatography when necessary

Peptide	Amino acid																	
	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr	Phe	Lys	AetCys	His	Arg	Trp
CB1T1Th1	1.0						1.0			1.0								
CB1T1Th2				1.1	0.9					0.8						1.0		+
CB1T1Th3													0.8	1.2				
T4Th1	1.1			0.9									2.0					
T4Th2				0.9		2.2							0.8					
T4Th3	1.1			0.9								0.7	1.1					
T4Th4		1.0	2.0								0.9							
T4C1				1.2								0.7						
T4C2	1.0				0.9						2.0		1.2					
T4C3		0.8	2.6		0.9					0.9	1.1	0.8						
T4C4												0.8					1.2	
T6Th1	1.1		0.9			1.1					0.8							
T6Th2			0.9	1.0						1.0								
T6Th3								1.1										0.9
T11Th1	1.0			2.2				0.8										
T11Th2				1.1				1.1					0.8					
T11Th3						1.2				0.8				1.2		0.8		
T13P1	2.2			0.8		1.1										0.9		
T13P2			1.1							1.0		0.7						0.9
T13C1	2.1			1.0		1.1						0.7				1.0		
T13C2			1.2							0.8								1.0
T17aTh1	0.9		1.1															
T17aTh2	1.2		1.0	1.1			0.8	0.9										
T17aTh3			1.1								0.9							+
CB2T1		1.0	0.9		1.1	1.0					0.9		1.1	1.0				
T18P1	1.0		1.2		0.9	1.2		0.9		0.8								
T18P2			1.1	1.0		1.1	0.8									0.9	1.0	
T20Th1			1.0	1.9	1.1									1.0				
T20Th2			2.9		1.0		1.0											

Table 4. Sequence determinations of the tryptic peptides (cf. Tables 2 and 3)

Some peptides were further degraded with thermolysin (Th), chymotrypsin (C) or pepsin (P). Fragment  $\alpha_A_2$ CB1 was used to establish the sequence of T1. Sequences were determined by Edman degradation using the direct phenylthiohydantoin method ( $\rightarrow$ ), by dansyl-Edman analysis ( $\leftarrow$ ), or by both methods ( $\leftrightarrow$ ). Carboxypeptidase C degradation was used on T20 ( $\leftarrow$ ). Residues indicated with  $\Rightarrow$  were identified after dansylation without being hydrolysed, which shows them to be present as free amino acids

Peptide	Sequence
T1	Ac-Met-Asp-Ile-Ala-Ile-Gln-His-Pro-Trp-Phe-Lys CB1 
T2	<u>Arg</u> $\Rightarrow$
T3	Thr-Leu-Gly-Pro-Phe-Tyr-Pro-Ser-Arg 
T4	Leu-Phe-Asp-Gln-Phe-Phe-Gly-Glu-Gly-Leu-Phe-Glu-Tyr-Asp- 
	-Leu-Leu-Pro-Phe-Leu-Ser-Ser-Thr-Ile-Ser-Pro-Tyr-Tyr-Arg 
T5	Gln-Ser-Leu-Phe-Arg 
T6	Thr-Val-Leu-Asp-Ser-Gly-Ile-Ser-Glu-Val-Arg 
T7-8	Ser-Asp-Arg-Asp-Lys 
T9	Phe-Val-Ile-Phe-Leu-Asp-Val-Lys 
T10	His-Phe-Ser-Pro-Glu-Asp-Leu-Thr-Val-Lys 
T11	Val-Gln-Glu-Asp-Phe-Val-Glu-Ile-His-Gly-Lys 
T12	His-Asn-Glu-Arg 
T13	Gln-Asp-Asp-His-Gly-Tyr-Ile-Ser-Arg 



Table 5. Amino-acid compositions of enzymic peptides overlapping the tryptic peptides of  $\alpha A_2$ . Data are given as molar ratios, without corrections for hydrolytic losses. These peptides were all purified by paper electrophoresis at pH 6.5 and, where necessary, at pH 3.8, followed by descending chromatography. The peptides are tabulated in the order in which they appear in the sequence alignment of Table 6

Peptide	Amino acid																	
	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr	Phe	Lys	AetCys	His	Arg	Trp
CB1Tm1	0.8			1.0	0.9		1.1			2.1			1.1	0.8		0.8	1.1	+
Th1		0.9											1.0	0.8			1.1	
C1		0.9									1.0			1.0			1.1	
C2			1.1		2.0	1.0					1.1	0.9	1.8				1.0	
Th2			1.0		1.1						0.9	1.1	0.8				1.0	
Th3			1.7	1.0	1.2					1.1		1.9					1.0	
C3			1.0	1.0							1.1	0.9	1.0				1.0	
Th4		0.9									1.0		1.0				1.1	
C4		0.8						1.2			1.2						0.9	
C5	2.7	0.9	2.8	1.2		1.1		2.1		1.0	0.9		1.0	1.1			3.3	
C6	2.7		2.7	1.1		1.1		0.8		1.0			1.0	0.9			2.0	
Th5	2.2		1.1					0.8					0.9	1.1			2.1	
Tm6-7	6.0	0.9	1.8	5.1	1.1	1.2		4.3*		1.6*	2.2		4.0	4.1		3.0	1.8	
Tm6-7P1	2.0		0.9										1.0	1.1			1.0	
Th6	0.9							1.0			1.1			1.0		1.1		
C7	1.2							0.8			1.1		1.1	1.0		0.9		
Tm6-7P2	1.1	0.7		2.0				2.0						1.0				
Tm6-7P3	0.9			2.2		1.2		1.0		0.9				1.1		1.7	1.1	
Th7	2.8			2.1		2.1				0.7		1.0		1.1		3.1	1.2	
Th8			1.0	1.0						0.8							1.2	
C8			0.9	1.1						0.9			1.1				1.0	
C9												0.9				0.7	2.1	
Th9			2.1								1.1				0.8			
C10			1.9								1.1				+			
Th10			2.0		2.1	1.9				1.1			0.9	1.0				
C11	1.2		2.9	1.1	1.7	3.0	1.1	1.0		1.0				1.0		0.9	1.0	
C12			4.5	2.3	3.0		2.0	1.1		1.1				1.1			1.1	

\* Low values due to incomplete hydrolysis of the Val-72-Ile-73 bond.

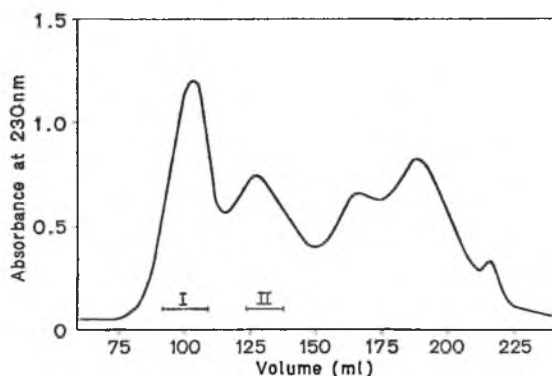


Fig. 6. Fractionation of a tryptic hydrolysate of maleylated aminoethyl-CB1. 1.5  $\mu$ mol digest was separated on Sephadex G-50 fine under the conditions given in Fig. 1. Peaks I and II contained Tm6-7 and Tm3, respectively which were further purified by rechromatography

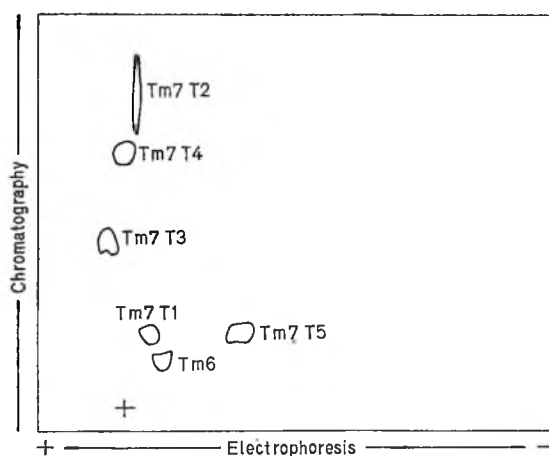


Fig. 7. Fingerprint of a tryptic digest of demaleylated Tm6-7. Conditions were the same as in Fig. 2

matographic behaviour. Tm6 is identical with T7, Tm7T1 is T8, Tm7T2 is T9, Tm7T3 is T10, Tm7T4 is T11 and Tm7T5 is T12. The sum of the sequences of these tryptic peptides agrees with the amino acid composition of Tm6-7.

These data together permit the following conclusions. The lysyl peptide T1 is followed in the sequence by T2, which is free arginine. The lysyl peptides T7-8, T9, T10 and T11 occur together in Tm6-7 and must be followed by the arginyl peptide T12. The other arginyl peptides, *i.e.* T3 through T7 and T13 through T16, as well as the aminoethyl-cysteinyl peptide T17a, must all be preceded by arginine.

#### Sequence Assembly

The evidence that CB1 is the N-terminal and CB2 the C-terminal CNBr fragment of the  $\alpha A_2$  chain

has already been discussed. Comparison of the fingerprints of tryptic digests of Aet- $\alpha A_2$ , Aet-CB1 and CB2 enabled to determine which tryptic peptides occurred in CB1 and CB2, respectively. CB1 was found to contain the peptides T1 (without acetyl-methionine) up to T17a, while the peptides T18, T19 and T20 were present in CB2.

Chymotryptic and thermolytic digests of Aet- $\alpha A_2$  and a peptic digest of Tm6-7 were used to isolate overlapping peptides that gave decisive information for the reconstruction of the arrangement of all tryptic peptides in both CNBr peptides. The amino acid analyses of those peptides which provided information on the alignment of tryptic peptides are given in Table 5. Table 6 shows then the way in which these overlapping peptides fix the order of the tryptic peptides, and thus establish the complete sequence of the  $\alpha A_2$  chain.

The information obtained from the overlap peptides is in agreement with the results of the tryptic digest of maleyl Aet-CB1, as discussed above.

**T1-T2-T3.** The N-terminal sequence of  $\alpha A_2$  was earlier reported from our laboratory to be Ac-Met-Asp-Ile-Ala [10]. The amino acid composition<sup>\*</sup> of T1 corresponds, apart from the methionyl residue, to the first ten residues of CB1 (Table 4). This places T1 at the N-terminus of the  $\alpha A_2$  chain. The T1-T2 linkage follows from the amino acid composition of CB1Tm1. The T1-T2-T3 junction was established by the sequence of peptides C1 and Th1.

**T3-T4-T5-T6.** The connection between T3 and T4 was evident from the structure of C2 and by the composition of Th2. The overlap T4-T5 was deduced from Th3. Th3 also confirms the C-terminal sequence of T4. The T5-T6 junction was arrived at by the sequence of Th4. Peptides C4 and C5 gave only limited information on the T5-T6 overlap as they only show that T6 is preceded by arginine.

**T6-T7-T8-T9-T10-T11-T12.** The order of T6, T7, T8 and T9 is firmly established by the sequence of Th5, as well as by the composition of C5, C6 and Tm6-7P1. Peptides T7 up to T12 form together the maleylated tryptic peptide Tm6-7. T12 must be located at its C-terminus, for it is the arginine-containing peptide. T9 and T10 are linked by the sequence of C7 and the composition of Th6, T10-T11 by the sequence of Tm6-7P2 and T11-T12 by Th7 and Tm6-7P3.

**T12-T13-T14-T15-T16-T17a-T17b.** The linkage of T12 and T13 was deduced from the sequence of Th7. The T13-T14 overlap is reached with Th8 as well as C8. The T14-T15-T16 junction was established after sequencing of C9 and by the occurrence in the tryptic digest of  $\alpha A_2$  of peptide T15-16. The tryptic digest also contains peptide T16-17a. The order of T17a-T17b was fixed by the sequence of C10 and the composition of Th9. This completes the ordering of the tryptic peptides in aminoethyl-CB1.

Table 6. Proposed primary structure of the  $\alpha A_2$  chain of bovine  $\alpha$ -crystallin

The overlapping peptides, determining the order of the tryptic peptides, are indicated. The amino acid compositions of these peptides are given in Table 5, but for T15-16 and T16-17a (Table 2) and CB2 (Table 1). Sequences were established by the dansyl-Edman method ( $\rightarrow$ ). The symbol  $-/-$  indicates that the concerning peptide was further digested with trypsin, and the resulting peptides identified by peptide mapping

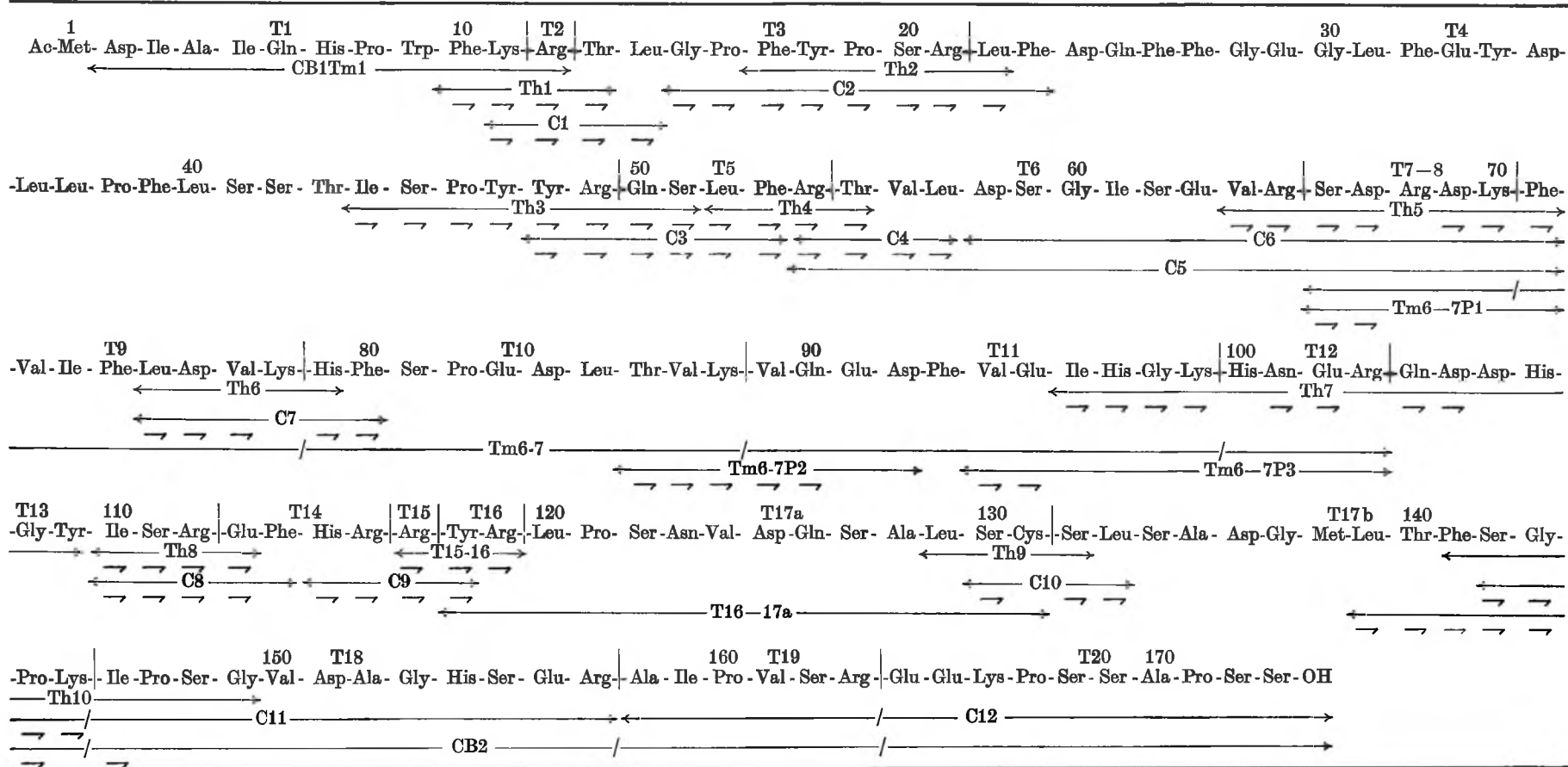


Table 7. Assignment of amide residues in the  $\alpha A_2$  chain

Methods of identification are: HVE, high-voltage electrophoresis, pH 6.5 (deduced charge in parentheses); PTH, direct identification of the phenylthiohydantoin derivative; EH, enzymatic hydrolysis

Residue	Peptide	Table	Identification	Conclusion
Asx-2	CB1T1Th1	3	HVE(-1)	Asp
Glx-6	CB1T1Th2	3	HVE(+1)	Gln
Asx-24	T4	2	PTH	Asp
Glx-25	T4	2	PTH	Gln
Glx-29	T4Th2	3	HVE(-1), PTH	Glu
Glx-33	T4Th3	3	HVE(-2), PTH	Glu
Asx-35	T4Th3	3	HVE(-2), PTH	Asp
Glx-50	T5	2	HVE(+1)	Gln
Asx-58	T6Th1	3	HVE(-1), PTH	Asp
Glx-63	T6Th2	3	HVE(-1), PTH	Glu
Asx-67	T7-8	2	HVE(0), PTH	Asp
Asx-69	T7-8	2	HVE(0), PTH	Asp
Asx-76	C7	5	HVE(+1)	Asp
Glx-83	T10	2	HVE(0), PTH	Glu
Asx-84	T10	2	HVE(0), PTH	Asp
Glx-90	T11Th1	3	HVE(-2, undegraded), PTH	Gln
Glx-91	T11Th1	3	HVE(-2, after 2 Edmans), PTH	Glu
Asx-92	T11Th1	3	HVE(-1, after 3 Edmans), PTH	Asp
Glx-95	T11Th2	3	HVE(-1), PTH	Glu
Asx-101	T12	2	HVE(+1), PTH, EH	Asn
Glx-102	T12	2	HVE(+1), PTH, EH	Glu
Glx-104	T13P1	3	HVE(-1) } N-terminus cyclizes	Gln
Asx-105	T13P1	3	HVE(-1) } remaining 2 residues	Asp
Asx-106	T13P1	3	HVE(-1) } must be acidic	Asp
Glx-113	T14	2	HVE(+1), PTH	Glu
Asx-123	T17aTh1	3	HVE(0), PTH	Asn
Asx-125	T17aTh2	3	HVE(-1), PTH, EH	Asp
Glx-126	T17aTh2	3	HVE(-1), PTH, EH	Gln
Asx-136	T17b	2	HVE(0), PTH	Asp
Asx-151	T18P1	3	HVE(-1)	Asp
Glx-156	T18P2	3	HVE(+1)	Glu
Glx-164	T20Th1	3	HVE(-1), PTH	Glu
Glx-165	T20Th1	3	HVE(-1), PTH	Glu

*T17b-T18-T19-T20.* CB2 yields four tryptic peptides, of which CB2T1 corresponds to the C-terminal portion of T17b. The first eight sequence steps of CB2 and the compositions of Th10 and C11 established the order T17b-T18. T20 must be the C-terminal peptide of the  $\alpha A_2$  chain, since it is the only tryptic peptide not terminated by lysine or arginine and its C-terminal residue, serine, corresponds to the known C-terminal residue of the  $\alpha A_2$  chain [2]. T19 and T20 occur together in C12, which places T19 penultimate in the alignment of the tryptic peptides of the  $\alpha A_2$  chain.

#### Amide Distribution

The results of the amide assignment are summarized in Table 7. Complete enzymic digestion was applied in only two cases (T12 and T17aTh2). Electrophoresis of T11Th1 after each Edman degradation step established the amide and acid groups in the sequence 90-91-92 as Gln-Glu-Asp, in agreement with the results of the phenylthiohydantoin determinations. Peptide T13P1, with the sequence Glx-Asx-

Asx-His-Gly, has a charge at pH 6.5 of -1, which implies the presence of two acidic residues. Since the N-terminal Glx was found to be able to cyclize to pyrrolidone carboxylic acid, it must be in the amide form, which leaves the acidic groups for the two Asx residues.

#### DISCUSSION

The data presented in this article have established the primary structure of the  $\alpha$ -crystallin  $A_2$  chain as given in Table 6, and show a total number of 173 amino acid residues. From these sequence data the molecular weight can be calculated to be 19832. This value is in good agreement with chemical and physicochemical measurements by various authors [2, 3, 8, 34].

Partial sequences of  $\alpha A_2$  published previously [6-9] are confirmed by the present results.

Comparison of the total amino acid composition derived from the proposed sequence with the amino acid analysis of  $\alpha A_2$  only reveals a significant discrepancy for glutamic acid. The value for this amino

acid found after amino acid analysis, is almost one residue too high. In fact recalculation of amino acid analysis data of  $\alpha A$  chains given by other authors [2, 34, 35] also shows a number of approximately 18 glutamyl residues per molecular weight of 19800. On the other hand, the number of glutamyl residues obtained by amino acid analysis of the CNBr peptides is in complete agreement with the sequence data. The Val-Ile bond in T9 is responsible for low values of valine and isoleucine after 24-h hydrolysis.

The nature of the peptide bonds cleaved by trypsin, chymotrypsin, thermolysin and pepsin, as can be seen in the Tables 4 and 6, is generally in accord with the known specificities of these enzymes. As expected, the Lys-Pro bond in T20 resists cleavage by trypsin. It is well known that acidic residues which are adjacent to a lysyl or an arginyl residue tend to reduce significantly the rate of hydrolysis by trypsin. Therefore the presence of Asp-67 adjacent to Arg-68 explains the very limited tryptic cleavage of the bond between T7 and T8 (or Tm6-Tm7). We did, however, not find any indication that Asp-69 hinders the cleavage behind Lys-70. The incomplete hydrolysis at the C-terminal side of Arg-117 and Arg-119 gave rise to two additional peptides: T15-16 and T16-17a, respectively.

Some chymotrypsin-like activity was apparently present in the trypsin preparations used. Most amino acid analyses of T4 showed non-integral values for tyrosine and arginine, due to partial cleavage between Tyr-47 and Tyr-48. This sets free the dipeptide Tyr-Arg, which is identical in composition to T16. A small amount of a peptide with the composition of T4C3 was often found in the tryptic digest as a result of an additional chymotrypsin-like cleavage at Phe-39. Likewise, hydrolysis at Phe-141 led in some cases to the appearance of the peptide Ser-Gly-Pro-Lys in the tryptic digest of  $\alpha A_2$ .

Thermolysin specifically cleaved N-terminal of isoleucine, valine, phenylalanine and leucine, except in the case of T17a and T20 where cleavage took place at the N-terminal side of serine. An unexpected cleavage by chymotrypsin has been encountered at the arginyl-alanyl bond (157-158) in a segment of the chain relatively devoid of sites susceptible to the action of chymotrypsin. Despite the presence of leucyl and phenylalanyl residues, parts of Tm6-7 were in our hands rather resistant to chymotryptic hydrolysis.

Peptides isolated by paper electrophoresis and chromatography proved to be as useful for direct Edman degradation as peptides isolated by column chromatography. As stressed by Hermodson *et al.* [36] quantitation of the more polar phenylthiohydantoyl residues even after silylation is difficult. Nevertheless quantitative analysis of these amino acid derivatives offers a possibility of calculating the

repetitive yield, which is particularly useful in the case of repeating sequences of the same amino acid. 0.2–0.4  $\mu\text{mol}$  peptide was sufficient for the direct Edman degradation of the tryptic peptides presented in Table 4. Abnormal behaviour of histidine during Edman degradation using volatile buffers has already been reported [37, 38]. We likewise experienced, that when degradation reached histidine in a sequence, not only this residue but also part of the next residue was released.

Tryptic digestion of the maleylated protein was very useful for the isolation of the cluster of lysyl peptides in the middle of the molecule. Partial maleylation of a few seryl residues was observed but did not diminish the usefulness of the method.

Robinson *et al.* [39] have indicated a correlation between the number of glutamyl and asparaginyl residues in a protein and its biological lifetime.  $\alpha A_2$  contains 4.6% amide residues which places it among the proteins with low amide content like collagen and histone IV. This would suggest a long half-life for the  $\alpha A_2$  chain, in agreement with the notion that the average turnover of lens proteins is very slow.

Recently the primary structure of bovine  $\gamma$ -crystallin (fraction II), another soluble lens protein, has been published by Croft [40]. Apparently no part of this sequence seems to be homologous to the  $\alpha A_2$  chain. On the other hand, structural homology between the  $\alpha A$  and  $\alpha B$  chains has been suggested since their N-terminal tryptic peptides show a striking resemblance [6, 34].

Our preliminary data on the amino acid compositions of the tryptic peptides of  $\alpha B_2$  demonstrate a homology of approximately 70% with the  $\alpha A_2$  chain (unpublished results).

It appears that the primary structure of  $\alpha A_2$  is highly preserved during evolution. This was concluded from comparative studies of soluble tryptic peptides of  $\alpha A_2$  chains from bovine, pig, dog and rabbit [41].

Although it is not possible to draw definite conclusions about the secondary and tertiary structure of  $\alpha A_2$  from these sequence studies, the distribution of hydrophobic and charged residues shows some interesting features. The N-terminal region (1–77) has a pronounced hydrophobic character, especially in T4 and T9. The subsequent region 78–119 is highly polar and has a distinct basic charge. The segment 120–144 is lacking any positive residue while the C-terminal part (145–173) is rather hydrophobic and rich in proline residues. It is tempting to speculate, that the hydrophobic N-terminal part of the sequence plays a special role in the aggregation behavior of the  $\alpha$ -crystallin molecule.

The work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) and by financial aid from the Netherlands Organization for Pure Research (Z.W.O.). We thank Mrs M. Versteeg and Mr G. Groenewoud for performing amino acid analyses.



## REFERENCES

1. Bloemendal, H., Berns, A. J. M., Zweers, A., Hoenders, H. J. & Benedetti, E. L. (1972) *Eur. J. Biochem.* **24**, 401–406.
2. Schoenmakers, J. G. G., Gerding, J. J. T. & Bloemendal, H. (1969) *Eur. J. Biochem.* **11**, 472–481.
3. Bloemendal, H., Rotmans-van Poppel, M. & van der Ouderaa, F. (1972) *FEBS Lett.* **28**, 81–85.
4. Bloemendal, H., Berns, A. J. M., van der Ouderaa, F. & de Jong, W. W. W. (1972) *Exp. Eye Res.* **14**, 80–81.
5. Delcour, J. & Papaconstantinou, J. (1972) *J. Biol. Chem.* **247**, 3289–3295.
6. Corran, P. H. & Waley, S. G. (1969) *Biochem. J.* **115**, 789–796.
7. Leon, A. E., Gerding, J. J. T., de Groot, K., Hoenders, H. J. & Bloemendal, H. (1971) *Int. J. Protein Res.* **3**, 19–24.
8. Corran, P. H. & Waley, S. G. (1971) *Biochem. J.* **124**, 61–67.
9. van der Ouderaa, F., de Jong, W. W. & Bloemendal, H. (1972) *FEBS Lett.* **28**, 77–80.
10. Hoenders, H. J., van Tol, J. & Bloemendal, H. (1968) *Biochim. Biophys. Acta*, **160**, 283–285.
11. Berns, A. J. M., van Kraaikamp, M., Bloemendal, H. & Lane, C. D. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 1606–1609.
12. Mathews, M. B., Osborn, M., Berns, A. J. M. & Bloemendal, H. (1972) *Nat. New Biol.* **236**, 5–7.
13. Berns, A. J. M., Schreurs, V. V. A. M., van Kraaikamp, M. W. G. & Bloemendal, H. (1973) *Eur. J. Biochem.* **33**, 551–557.
14. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
15. Schoenmakers, J. G. G., Matze, R., van Poppel, M. & Bloemendal, H. (1969) *Int. J. Protein Res.* **1**, 19–27.
16. Raftery, M. A. & Cole, R. D. (1966) *J. Biol. Chem.* **241**, 3457–3461.
17. Butler, P. J. G., Harris, J. H., Hartley, B. S. & Lebermann, R. (1969) *Biochem. J.* **112**, 679–689.
18. Grosclaude, F., Mercier, J. C. & Ribadeau-Dumas, B. (1970) *Eur. J. Biochem.* **14**, 98–107.
19. Gross, E. (1967) *Methods Enzymol.* **11**, 238.
20. Tschesche, H. & Kupfer, S. (1972) *Eur. J. Biochem.* **26**, 33–36.
21. Waley, S. G. & Watson, J. (1953) *Biochem. J.* **55**, 328–337.
22. Easley, C. W., Zegers, B. J. M. & de Vylder, M. (1969) *Biochim. Biophys. Acta*, **175**, 211–213.
23. Jones, R. T. (1970) in *Methods of Biochemical Analysis* (ed. D. Glick) vol. 18, p. 205–258, John Wiley and Sons, New York.
24. Reindel, F. & Hoppe, W. (1954) *Chem. Ber.* **87**, 1103–1107.
25. Edelhoch, H. (1967) *Biochemistry*, **6**, 1948–1954.
26. Niall, H. D. & Potts, J. T. (1970) in *Peptides, Chemistry and Biochemistry* (Weinstein, B. and Lande, S. eds) p. 215–241, Dekker, New York.
27. Braunitzer, G., Schrank, B., Ruhfus, A., Petersen, S. & Petersen, U. (1971) *Hoppe-Seyler's Z. Physiol. Chem.* **352**, 1730–1732.
28. Gray, W. R. & Smith, J. F. (1970) *Anal. Biochem.* **33**, 36–42.
29. Pataki, G. (1966) *Dünnschicht-Chromatographie in der Aminosäure und Peptid Chemie* pp. 173–190, Walter de Gruyter & Co., Berlin.
30. Jeppson, J. O. & Sjöquist, J. (1967) *Anal. Biochem.* **18**, 264–269.
31. Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597–5607.
32. Woods, K. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta*, **133**, 369–370.
33. Offord, R. E. (1966) *Nature (Lond.)* **211**, 591–593.
34. Augusteyn, R. C. & Spector, A. (1971) *Biochem. J.* **124**, 345–355.
35. Palmer, W. G. & Papaconstantinou, J. (1969) *Proc. Natl. Acad. Sci. U. S. A.* **64**, 404–410.
36. Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H. & Walsh, K. A. (1972) *Biochemistry*, **11**, 4493–4502.
37. Blömbäck, G., Blömbäck, M., Hessel, B. & Wanaga, S. (1967) *Nature (Lond.)* **215**, 1445–1448.
38. Schroeder, W. A. (1967) *Methods Enzymol.* **11**, 445–461.
39. Robinson, A. B., McKerrow, J. H. & Cary, P. (1970) *Proc. Natl. Acad. Sci. U. S. A.* **66**, 753–757.
40. Croft, L. R. (1972) *Biochem. J.* **128**, 961–970.
41. de Jong, W. W., van Amelsvoort, J. M., van der Ouderaa, F. J. & Bloemendal, H. (1973) *Genetics*, **74**, a60.

F. J. van der Ouderaa, W. W. de Jong and H. Bloemendal, Laboratorium voor Biochemie der Universiteit, Geert Groteplein Noord 21, Nijmegen, The Netherlands