Biochimica et Biophysica Acta, 491 (1977) 573–580
© Elsevier/North-Holland Biomedical Press

BBA 37626

PRIMARY STRUCTURES OF α-CRYSTALLIN A CHAINS OF ELEPHANT, WHALE, HYRAX AND RHINOCEROS

WILFRIED W. DE JONG, EUGÉNIE C. NUY-TERWINDT and MARLIES VERSTEEG
Department of Biochemistry, University of Nijmegen, Nijmegen (The Netherlands)
(Received November 11th, 1976)

SUMMARY

As part of a study of the evolutionary development of the eye lens protein α-crystallin the 173-residue A chain of this protein has been studied in elephant, whale, hyrax and rhinoceros. The primary structures were inferred mainly from amino acid compositions of peptides obtained by enzymic digestions and CNBr cleavage. The positions of substitutions, as compared to the known bovine A chain, were confirmed by Edman degradation. In accordance with the previously observed slow rate of evolution of the A chain only a small number of substitutions was found among these species. Elephant and hyrax share a number of unique substitutions, strongly indicating a common ancestry of these two species within the mammalian class.

INTRODUCTION

One of the most promising applications of comparative protein-sequence studies is the reconstruction of phylogenetic relationships between organisms [1, 2]. The eye-lens protein α-crystallin shows a relatively slow rate of evolution [3], and might therefore be well-suited to study the relationships between higher vertebrate taxa. To this end we studied already the amino acid sequences of the α-crystallin A chains of eleven mammalian species [4–7], belonging to seven orders. These mostly represented easily obtainable and commonly studied species: ox, pig, horse, dog, cat, rabbit, rat, rhesus monkey, man, kangaroo and opossum.

To further unravel the phylogenetic history of the mammalian orders it is necessary to study representatives of additional orders. We succeeded in obtaining eye-lenses from elephant, whale, hyrax and rhinoceros, belonging to the orders *Proboscidea*, *Cetacea*, *Hyracoidea* and the suborder *Ceratomorpha* of the *Perissodactyla*, respectively. The primary structures of the α-crystallin A chains of these species are reported here.

MATERIALS AND METHODS

Frozen lenses or whole eyes were shipped to Nijmegen. The species were: African elephant (*Loxodonta africana*; 11 lenses, average wet weight per lens 1.4 g),

minke whale or lesser rorqual (*Balaenoptera acutorostrata*; five lenses of 3.3 g), Cape hyrax or rock dassie (*Procavia capensis*; 30 lenses of 0.15 g), square-lipped rhinoceros (*Ceratotherium simum*; four lenses of 0.8 g). All animals were adult or sub-adult.

The lenses of each species were pooled before isolating the water-soluble proteins. All subsequent procedures have been described in more detail previously [4,5,7]. α-Crystallin was isolated by gel filtration on Sephadex G-200 and the A and B chains were separated by column chromatography on carboxymethyl-cellulose in the presence of 7 M urea. Acrylamide gel electrophoresis of α-crystallin and its chains was carried out in Tris/glycine buffer, pH 8.5, containing 6 M urea [8].

The isolated A chains were aminoethylated before digestion with trypsin. In some instances the A chains or tryptic peptides were digested with thermolysin, chymotrypsin, pepsin or staphylococcal protease. Soluble peptides were separated by high voltage electrophoresis (pH 6.5) on paper, followed by descending chromatography. After staining with fluorescamine, peptides were eluted from the paper, analysed for amino acid composition and, when required, used for dansyl-Edman degradation. Amide assignment was based on the electrophoretic mobility of peptides at pH 6.5. CNBr cleavage and maleylation followed by tryptic digestion were necessary to elucidate the substitutions in the A chains of elephant and hyrax.

RESULTS

α-Crystallin could be isolated by a single step of gel filtration. It formed 49%, by weight, of the water-soluble lens proteins in the elephant, 53% in the whale, 35% in the hyrax and 23% in the rhinoceros. The subunit patterns of these α-crystallins, as revealed by acrylamide gel electrophoresis, are shown in Fig. 1. The A_2 chains of whale and rhinoceros move in the same position as observed in all hitherto studied placental mammals [5]. By contrast the A_2 chains of elephant and hyrax are both more negatively charged.

Apart from the α-crystallin chains A_2 and B_2, and their deamidation products A_1 and B_1, which are the main components of calf α-crystallin, additional chains appear upon ageing in the bovine lens, and are due to degradation and further deamidation of A and B chains [9]. The pattern of the ageing chains is quite similar in rhinoceros and horse, which belong to the same order. The degree of deamidation and degradation of the primary gene products A_2 and B_2 is remarkably small in elephant and hyrax.

The patterns of the tryptic peptide maps of the aminoethylated A_2 chains of the four species are compared in Fig. 2 with the known bovine pattern [4]. It should be mentioned that in no case was free lysine observed on the peptide maps, thus excluding the possibility that in any species one of the two residues of arginine, which are normally found in the tryptic digest of the bovine A_2 chain, has been substituted by lysine.

The tryptic peptides insoluble at pH 6.5 were separated by gel filtration and resulted in the isolation of T4 and T9 in whale and rhinoceros, as in the ox [4], and T4 only in elephant and hyrax. Peptide maps of the thermolytic digests of T4 were indistinguishable in all four species, and identical to the previously studied A chains [5]. Tryptic peptides homologous to the bovine ones and accounting for the complete A_2 chains of the four species, apart from T7–8 and T9 of elephant and hyrax, were

isolated and analysed*. Additional digestions with other enzymes and cleavage with CNBr were performed where necessary to localize substitutions. The identified substitutions are summarized in Table I.

TABLE I

Identified substitutions in the A chains of α-crystallin of elephant, whale, hyrax and rhinoceros, as compared to the bovine A chain. In position 55 of the hyrax A chain equal amounts of alanine and threonine are present.

Position	3	4	13	55	61	70	72	74	90	91	127	142	146	147	150	153
Ox	Ile	Ala	Thr	Thr	Ile	Lys	Val	Phe	Gln	Glu	Ser	Ser	Ile	Pro	Val	Gly
Elephant	Val	Thr	Ala	Thr	Ile	Gln	Val	Leu	Gln	Asp	Ser	Cys	Ile	Gln	Met	Ser
Whale	Ile	Ala	Ala	Thr	Ile	Lys	Val	Phe	Gln	Glu	Ser	Ser	Val	Pro	Met	Gly
Hyrax	Val	Thr	Ala	Thr Ala	Ile	Gln	Leu	Leu	Leu	Asp	Ser	Cys	Val	Gln	Met	Ser
Rhinoceros	Ile	Ala	Thr	Ser	Val	Lys	Val	Phe	Gln	Glu	Thr	Ser	Ile	Pro	Met	Gly

Elephant

Differences in amino acid composition, as compared to the corresponding bovine peptides, were observed in T1, T3, T11, T17b and T18. In addition a tripeptide Gly-Pro-Lys was detected on the peptide map, and T17b was positively charged instead of neutral as in the ox (Fig. 2). The substitutions in these peptides were identified as indicated in Fig. 3. The substitution 142 Ser → Cys explains the positive charge of T17b after aminoethylation of the A_2 chain. Tryptic cleavage behind 142 aminoethylcysteine is apparently inefficient, but is responsible for the small amount of tripeptide Gly-Pro-Lys on the fingerprint.

Peptides T7–8 and T9 were neither found on the peptide map, nor by gel filtration of the insoluble peptides This region of the elephant A_2 chain was explored by making fingerprints of chymotryptic, thermolytic and staphylococcal protease digests of the chain. The informative peptides C4, Th1 and SP1 were isolated, analysed and partially sequenced (Fig. 3). This demonstrated the substitution 70 Lys → Glx, which must be glutamine on the basis of the electrophoretic mobility of peptide C4.

* Supplementary data to this article are deposited with, and can be obtained from, Elsevier Scientific Publishing Company, BBA Data Deposition, P.O. Box 1527, Amsterdam, The Netherlands. Reference should be made to No. BBA/DD/056/37626/491(1977)575.

The supplementary information contains: (1) Amino acid compositions of all tryptic peptides of the aminoethylated A chains of α-crystallin from elephant, whale, hyrax and rhinoceros. (2) Amino acid compositions of the thermolytic peptides obtained from T4 of elephant, whale, hyrax and rhinoceros. (3) Amino acid compositions of CB2, CB3 and CB1Tm6 (obtained by tryptic digestion of maleylated CB1) of elephant and hyrax. (4) Amino acid compositions of some chymotryptic, thermolytic, staphylococcal protease, peptic and tryptic peptides obtained from total A chains or by secondary digestion of tryptic or CNBr peptides, used to deduce the substitutions in elephant and hyrax A chains. (5) Tracings of peptide maps of chymotryptic, thermolytic and staphylococcal protease digests of the elephant A chain. (6) Tracings of the paper electrophoretic separations of the thermolytic peptides obtained from elephant T1, and of the peptic peptides obtained from hyrax T13 + T18.

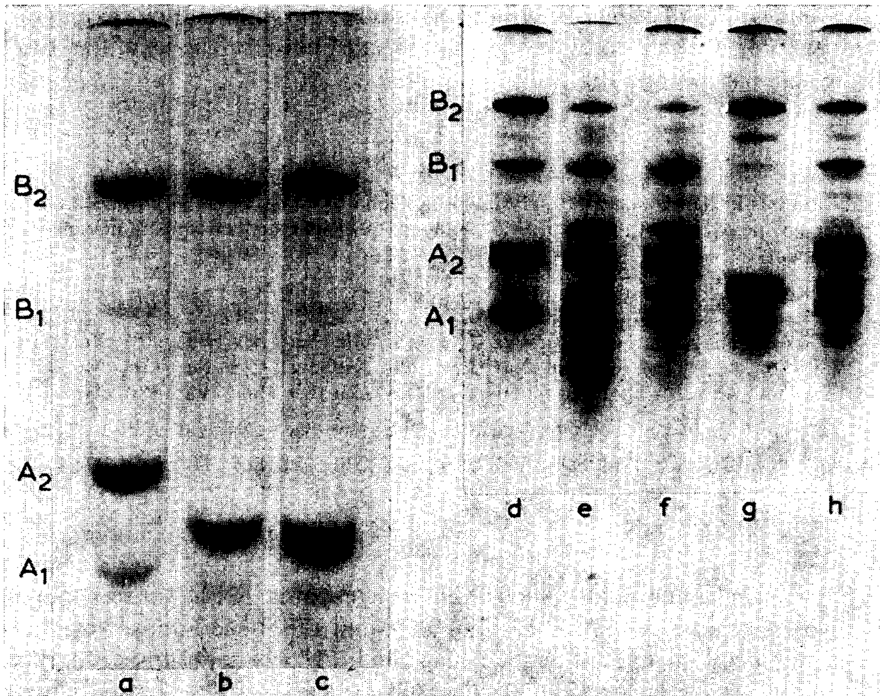

Fig. 1. Acrylamide gel electrophoresis in Tris/glycine buffer, pH 8.5, containing 6 M urea. Samples are α-crystallin from (a) ox, (b) elephant, (c) hyrax, (d) ox, (e) horse, (f) rhinoceros, (g) elephant, and (h) whale.

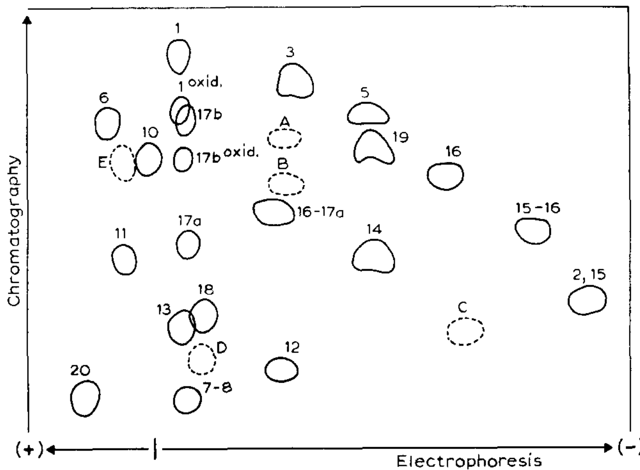

Fig. 2. Composite peptide map of the tryptic digests of the aminoethylated A_2 chains of ox, elephant, whale, hyrax and rhinoceros. Electrophoresis at pH 6.5 was followed by descending chromatography. Solid circles indicate the pattern observed from the bovine chain [4]. The tryptic peptides are numbered in the order in which they occur in the A chain (see Fig. 3). Peptides T4 and T9 are insoluble at pH 6.5 and are isolated by gel filtration over Sephadex G-50 [4]. A and B indicate the positions of T17b and $T17b^{oxid.}$, respectively, in elephant and hyrax; C represents the tripeptide Gly-Pro-Lys which is only found in elephant and hyrax; D corresponds to the oxidized form of T18 which is found in elephant, whale, hyrax and rhinoceros; E is the position of T11 in the hyrax. Peptide T7-8 is absent on the elephant and hyrax peptide map.

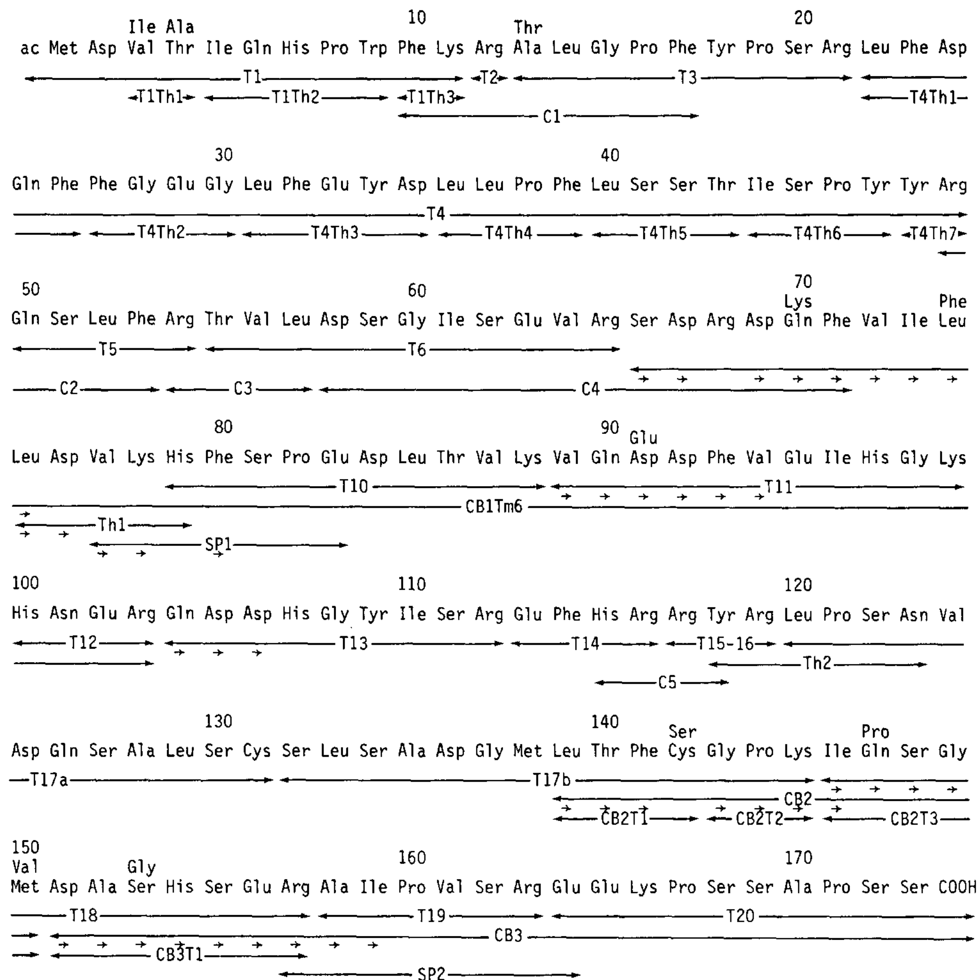

Fig. 3. Proposed amino acid sequence of the A_2 chain of elephant α-crystallin. The figure summarizes which peptides were isolated for amino acid analysis, and how residues were placed by Edman degradation ($\rightarrow$). Residues which are different in the bovine A_2 chain are given above the elephant sequence. Tryptic (T), chymotryptic (C), thermolytic (Th), staphylococcal protease (SP) and CNBr (CB) peptides were obtained from the complete chain. Peptides T1, T4, CB2 and CB3 were further degraded with thermolysin or trypsin. Fragment CB1 was digested with trypsin after maleylation to obtain peptide CB1Tm6 [4]. The patterns of peptide maps from which all the indicated peptides are isolated, and the amino acid analyses of these peptides are contained in the deposited supplementary information.

In addition the large peptide comprising residues 66–103 was isolated by gel filtration after tryptic digestion of the maleylated fragment CB1 of the A_2 chain. Determination of the N-terminal sequence of this peptide revealed the still missing residues from position 72 to 74 (Fig. 3), including the substitution 74 Phe $\rightarrow$ Leu.

Whale

Tryptic peptides corresponding to all peptides of the bovine A_2 chain were

isolated. Differences in composition were only present in T3 and T18. N-Terminal determination of T3 revealed the substitution 13 Thr → Ala, and eight steps of Edman degradation on T18 the substitutions 146 Ile → Val and 150 Val → Met.

Hyrax

The fingerprint of hyrax aminoethylated A_2 chain differed from that of the elephant in the displacement of T11 (Fig. 2). Differences in amino acid composition, as compared to the homologous bovine peptides, were observed in T1, T3, T6, T11, T17b and T18. The substitutions were, as in the elephant, located by thermolytic digestion of T1, N-terminal determination of T3, three steps of Edman degradation on T11, and deduced from the partial tryptic cleavage after 142 aminoethylcysteine. T18 was isolated contaminated with T13. After peptic digestion of the mixture [4] the two peptides derived from T18 were purified and partially sequenced, showing the substitutions 146 Ile → Val, 147 Pro → Gln, 150 Val → Met and 153 Gly → Ser. The substitution 150 Val → Met was confirmed by CNBr cleavage behind this residue, and 142 Ser → Cys by the tryptic peptides derived from CB2 fragment (as in the elephant).

Three steps of Edman degradation on T6 showed the presence of both threonine and alanine in position 55, in agreement with the amino acid composition, which yielded half integral amounts for each alanine and threonine. The same non-integral values for alanine and threonine were found in hyrax C3.

Like in the elephant, hyrax T7–8 and T9 could not be isolated. Isolation and ten steps of Edman degradation on the tryptic peptide Tm6 of the maleylated CB1 fragment of the A_2 chain, corresponding to residues 66–103, revealed the substitutions 70 Lys → Gln, 72 Val → Leu and 74 Phe → Leu.

Isolation and amino acid analyses of peptides corresponding to elephant C4 and Th1 further confirmed the correctness of the proposed sequence of the hyrax A_2 chain.

Rhinoceros

Differences in amino acid composition, as compared to the corresponding bovine peptides, were only detected in T6, T17a and T18. Eight steps of Edman degradation on T6 revealed the substitutions 55 Thr → Ser and 61 Ile → Val, eight steps on T17a gave 127 Ser → Thr, and five steps on T18 gave 150 Val → Met. Because rhinoceros differed from horse A chain in the composition of peptides T3 and T17b, five steps of Edman degradation were performed on T3 and ten steps on T17b. In both cases the rhinoceros sequence was found to be identical to the bovine one.

DISCUSSION

The main purpose of the present work was to acquire additional mammalian A chain sequences to be used in the reconstruction of the phylogeny of mammalian orders. The actual construction of phylogenies from the now available fifteen A chain sequences requires complicated data-handling and will be dealt with elsewhere. However, at first glance some interesting features of the sequences presented here are evident.

Most remarkable is the correspondence between hyrax and elephant A chains. They distinguish themselves among all studied placental species by their electrophoretically faster moving α-crystallin A chain, due to the substitution 70 Lys →Gln. They share another two substitutions which are not found in the other species: 74 Phe → Leu and 153 Gly → Ser. The substitution 142 Ser → Cys is apart from hyrax and elephant only found in human A chain [10]. This number of unique substitutions is very suggestive of a common origin of the orders *Hyracoidea* and *Proboscidea*, an opinion held by some taxonomists [11], but denied by others [12].

Incidentally hyrax, elephant and human α-crystallin are all distinguished by the near absence of the A_1 chain. One wonders whether the presence of a second cysteine in the A chain, in position 142, somehow interferes with the deamidation process which is supposed to lead to the formation of A_1 from A_2 [13].

The rhinoceros shares only one unique substitution, 127 Ser → Thr, with the previously investigated perissodactyl, the horse. The substitution 55 Thr → Ser is unique for rhinoceros; 61 Ile → Val has only been found in pig, and is probably due to a parallel substitution; 150 Val → Met is found in several other mammals. The three substitutions found in whale A chain occur in different combinations in various mammals, and do not suggest any obvious relationship to other orders.

The presence of both alanine and threonine in position 55 of the hyrax A chain could be explained by the occurrence of two loci for the A chain, differing in the codon for residue 55, or by the presence of a mutant allele in the population at a frequency of about 50%. Although the latter explanation is somewhat more likely, a definite answer can only be provided by studying individual hyraxes for the nature of residue 55 in their A chains.

ACKNOWLEDGEMENTS

We are grateful to Drs. U. de V. Pienaar, V. de Vos and P. van Wyk, Kruger National Park, for providing elephant and hyrax lenses; to the Natal Parks, Game and Fish Preservation Board and Dr. M. E. Keep, Hluhluwe Game Reserve, South Africa, for rhinoceros lenses; and to Drs. Å. Jonsgård, Oslo, and I. Christensen, Bergen, for whale lenses. This work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) and with financial aid from the Netherlands Organization for Pure Research (Z.W.O.).

REFERENCES

1 Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure, Vol. 5, National Biomedical Research Foundation, Silver Spring, Md.
2 Peacock, D. and Boulter, D. (1975) J. Mol. Biol. 95, 513–527
3 De Jong, W. W., Van Amelsvoort, J. M., Van der Ouderaa, F. J. and Bloemendal, H. (1973) Nat. New Biol. 246, 233–236
4 Van der Ouderaa, F. J., De Jong, W. W. and Bloemendal, H. (1973) Eur. J. Biochem. 39, 207–222
5 De Jong, W. W., Van der Ouderaa, F. J., Versteeg, M., Groenewoud, G., Van Amelsvoort, J. M. and Bloemendal, H. (1975) Eur. J. Biochem. 53, 237–242
6 De Jong, W. W., Terwindt, E. C. and Bloemendal, H. (1975) FEBS Lett. 58, 310–313
7 De Jong, W. W. and Terwindt, E. C. (1976) Eur. J. Biochem. 67, 503–510
8 Bloemendal, H. (1967) in Electrophoresis. Theory, Methods and Applications (Bier, M., ed.), Vol. 2, p. 379, Academic Press, New York

9 Van Kleef, F. S. M., De Jong, W. W. and Hoenders, H. J. (1975) Nature 258, 264–266
10 Kramps, J. A. (1977) Ph.D. thesis, University of Nijmegen
11 Thenius, E. (1969) Phylogenie der Mammalia, Walter de Gruyter, Berlin
12 McKenna, M. C. (1975) in Phylogeny of the Primates (Luckett, W. P. and Szalay, F. S., eds.), pp. 21–46, Plenum, New York
13 Bloemendal, H., Berns, A. J. M., Van der Ouderaa, F. J. and De Jong, W. W. (1972) Exp. Eye Res. 14, 80–81