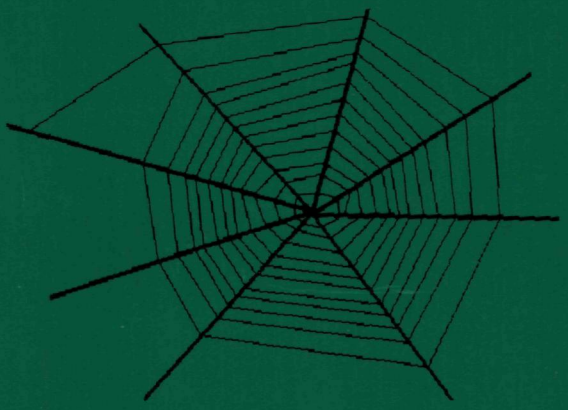
MOLECULAR EVOLUTION OF EYE LENS PROTEINS: CRYPTIC STRUCTURES AND FUNCTIONS



Wiljan Hendriks

MOLECULAR EVOLUTION OF EYE LENS PROTEINS: CRYPTIC STRUCTURES AND FUNCTIONS

MOLECULAR EVOLUTION OF EYE LENS PROTEINS: CRYPTIC STRUCTURES AND FUNCTIONS

Een wetenschappelijke proeve op het gebied van de
WISKUNDE EN NATUURWETENSCHAPPEN

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit te Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op
donderdag 16 februari 1989
des namiddags te 3.30 uur

door

WILHELMUS JOHANNES ANTONIUS JOSEPHUS HENDRIKS

geboren op 7 oktober 1959
te Haarlemmerliede en Spaarnwoude

Promotor: Prof.Dr. H. Bloemendal

Co-referent: Dr. W.W. de Jong

The investigations described in this thesis were carried out under the auspices of the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO) and performed under the direction of Dr. W.W. de Jong and Prof. Dr. H. Bloemendal at the Department of Biochemistry, Faculty of Science, University of Nijmegen, Nijmegen, The Netherlands.
This thesis was supported by Pharmacia Nederland B.V..

Graag wil ik op deze plaats iedereen bedanken die heeft bijgedragen aan het tot stand komen van dit proefschrift. Zonder iemand te kort te willen doen, wil ik toch enkelen met name noemen.

Allereerst de "primstrukkers", en niet alleen vanwege de lekkere werksfeer. De bijdragen van medepromovendi John, Christien en Jack gaan verder dan de auteurslijst bij de diverse hoofdstukken doet vermoeden. Marlies verzorgde vakkundig de aminozuuranalyses, en Willeke en Gerrit zorgden voor de talloze kleine dingen die 't doen, die 't doen. De nieuwe generatie "primstrukkers", Karin, Patricia, Perry en Ed, dank ik voor het feit dat ze mij, ondanks mijn "stressie toestand" van de laatste tijd, niet uit het lab hebben gegooid.

Ook wil ik, in chronologische volgorde, de doctoraalstudenten die mij de afgelopen jaren hebben begeleid, bedanken: Wouter Huizinga, Jo Sanders, Hans Weetink, Alex Wetzels, Matthijs Blankesteijn, Hans Polderman en Bernhard Kleter.

Jack Schalken, Geert van Rens, Piet van den Oetelaar, Winand Habets en Martin Salden bedank ik voor hun adviezen bij het opzetten en uitvoeren van diverse experimenten. Els, bedankt voor de essentiële koppen koffie en correspondentie. Alle overige (ex-) medewerkers van de afdeling Biochemie wil ik bedanken voor het prima zwoegklimaat.

Tevens wil ik de crystalline-groep van de afdeling Moleculaire Biologie bedanken voor de stimulerende contacten. Ook de vruchtbare samenwerking met Frans Ramaekers en Loe de Leij heb ik zeer gewaardeerd.

I would like to thank Eviatar Nevo for his "blind" contribution, Mladen Globulic, Filippe Figueroa, and Jan Klein for their hospitality and co-operation, and Mike Bibby, Christine Slingsby, and Tom Blundell for useful discussions and fruitful collaboration.

Natuurlijk wil ik ook mijn familie en vrienden bedanken. Zij zorgden ervoor dat ik me er steeds van bewust bleef dat niet alleen de lens de mens de moeite waard maakt.

Bedankt Wilma, niet alleen voor het wachten als het weer "even" uitliep, voor het overnemen van diverse huiselijke aangelegenheden als er "toch echt" moest worden doorgewerkt en voor je geduld als ik weer eens "enigszins gespannen" overkwam, maar voor heel veel meer!

In gedachtenis aan Richard

Voor mijn ouders

Voor Wilma

CONTENTS

CHAPTER 1

INTRODUCTION

General Introduction

The vertebrate eye lens is a crystal-clear, biconcave organ that has the ability to focus the incident light on the retina. Its most conspicuous morphological features are transparency, lack of direct blood supply and innervation, and continuous growth throughout life by forming new fiber cells at the outer part of the lens without ever replacing the older, inner, cells. As a consequence, the cells in the nucleus of the lens still contain proteins that were synthesized during the embryonic stage of the animal. To warrant transparency, the light-scattering organelles, including the nuclei, are broken down in the later stages of differentiation. The lens is therefore an attractive system for the study of fundamental biological processes like growth, differentiation, and aging (for reviews see 1-4).

Lens cells have a very high protein concentration, varying between 20% in the soft lenses of birds, and 50% in the very hard nuclei of fish lenses. This high protein content is necessary for lens transparency and for generating a refractive index gradient for proper light-focussing (5, 6). A small number of different polypeptides, making up more than 90% of the total lens protein, accounts for this high amount of protein. These very abundant, water-soluble proteins in the lens are called crystallins. Based on several criteria, four major crystallin classes can be distinghuished (7). α- and β-crystallins are common to all vertebrate species. Birds and reptiles are the only vertebrates which do not seem to have γ-crystallin; instead, they express δ-crystallin. Apart from these major crystallin species, some additional unrelated crystallins with a more restricted occurrence are found in different vertebrate groups (8-13).

Despite its similar function in all vertebrate eyes, the lens displays a huge variation in protein composition. Crystallins not only vary between species but are also differentially expressed during lens development (2). And, in addition, the lens cells increase the heterogeneity of their crystallin composition by an array of mechanisms at the transcriptional and translational level (7, 14). On the other hand, comparative sequence analysis has revealed that the investigated crystallins all evolve at very slow rates (10, 15). According to current evolutionary thinking (16-18) this implies that the structure of the crystallins must be very important with respect to their function. However, no common characteristic structural properties for the

different crystallins have yet emerged, although thermodynamic stability would seem to be an obvious advantage for proteins that have to survive in the long-living lens cells under relatively dehydrated conditions (5).

A type of eye similar to the ones of present-day vertebrates must have been present in the last common ancestor of all living vertebrates, approximately 400-500 million years ago. Evolution seldom proceeds by the de novo invention of new structures but rather elaborates them using pre-existing material of different origin and function, selected and modified for a new role (19). In the last eight years sequence analysis of the different crystallin classes has indeed provided insight into the thrilling evolutionary origin of the abundant lens proteins (for an overview see 20). This will be outlined below.

α-Crystallin

α-Crystallin has been observed in almost all vertebrate lenses; only some bony fishes seem to lack this protein aggregate that can reach levels of up to 50% of total lens protein in certain mammalian species (1). The 20 kDa subunits of this large molecule (300 to 1100 kDa depending on the species and the isolation procedures; e.g. 21, 22) are produced by a family of only two genes: αA and αB (7). Nevertheless, post-translational modifications (23) and in some cases alternative splicing (24) render the α-crystallin aggregate to be comprised of a heterogeneous array of sbunits. Structural studies are therefore dealing with a rather complex protein aggregate.

α-Crystallin subunits mainly contain β-pleated sheets, and almost no α-helices are detectable (25). The tertiary structure is not known but several hypothetical models have been suggested (26-28). The quarternary structure of α-crystallin has been examined exhaustively but still remains a matter of continuing debate. Several different models have been proposed. Some are based on the equivalence of subunit positions in the native α-crystallin aggregate (29, 30), others distinghuish different layers in the protein, which implies that subunits do not occupy equivalent sites (31, 32).

Protein sequences of α-crystallin subunits have been comprehensively studied throughout the vertebrates, serving as yardsticks for phylogenetic relationships (15, 33, 34). These data led to the discovery of the first known relationship between a crystallin and a non-lens-specific protein family. α-Crystallin turned out to be derived from the ubiquitous small heat-shock proteins (HSPs) by gene duplication and subsequent divergence, and

both are related to a major schistosome egg antigen (34-36). It has been found that the small HSPs in Chlamydomonas chloroplasts protect the photosynthetic machinery against light-induced damage (37). Therefore, it is tempting to speculate that α-crystallin was recruited as a lens structural protein because it shares thermodynamic stability with the small HSPs (28) and has a similar protective role against the many forms of stress during the lifetime of the lens (38).

Other Major Crystallins

β-Crystallins, like α-crystallins, are multimeric proteins. The subunits range in size from 20 to 30 kDa, and, with the exception of βs (renamed γs), form dimers or mixed aggregates of 50-200 kDa. γ-crystallins are exclusively monomeric, about 21 kDa in size (for review see 1-4). It has been shown that β- and γ-crystallins are members of the same protein superfamily (39), with γs being an old offshoot of the γ-crystallin branch (40). Based on structural criteria a conspicious similarity with a bacterial spore coat protein has been observed (41). They all display a fourfold internally repeated sequence, corresponding with four similar structural motifs, which are arranged into two compact domains (5, 7, 42). The γ-crystallins have a high proportion of cysteine and aromatic residues, which have been proposed to play an important role in protecting the lens proteins against the deleterious effects of ultraviolet radiation and other agents leading to the generation of free radicals (43).

Enzyme-related Crystallins

Interestingly, the other crystallins thusfar investigated all turn out to be related, or are even identical, to normal cellular enzymes (13, 20, 44). δ-Crystallin, the major protein in the embryonic lens of birds and reptiles, shares clear homology with an enzyme called argininosuccinate lyase (ASL) (45, 46). In the chicken genome there are two δ-crystallin genes (7). The δ1-crystallin gene is predominantly expressed in lens tissue but its product has no ASL activity, probably due to a mutation near the substrate binding-site of the enzyme (47). The δ2 gene is only weakly expressed in the chicken lens and probably, being the ortholog of the human ASL gene, fulfils the needs for active ASL elsewhere in the organism. Quite opposite, both δ1 and δ2 gene transcripts are found in large amounts in the duck lens, resul-

ting in an extremely high ASL activity for duck lens homogenates (47). The δ2-crystallin/ASL gene in duck therefore encodes a protein with two different functions. In non-lens tissues the normal "housekeeping" enzyme is produced, whereas in the lens large amounts of a structural protein is generated.

This phenomenon of "gene sharing" (47) is also displayed by ε- and τ-crystallin. ε-Crystallin, found in many bird and crocodile lenses (10), was the first lens protein for which homology with an enzyme was detected. Duck lens ε-crystallin was found to be almost indistinguishable from duck heart lactate dehydrogenase (LDH) B4 (48). Subsequent studies have shown that a single-copy Ldh-B gene in duck is heavily overexpressed in lens tissue, giving rise to ε-crystallin that differs from the primary gene product due to posttranslational modifications (49). In the chicken genome there is also a single-copy Ldh-B gene (49) but this does not exhibit the copious expression in lens cells since no ε-crystallin can be detected in the chicken lens (10, 48). The fact that LDH-B/ε-crystallin is subject to posttranslational modifications in the lens environment does not seem to have much influence on the specific activity of this glycolytic enzyme (48). Already in the early sixties it was reported that LDH-B in birds and higher reptiles exhibits an increased thermal stability (50), perhaps explaining why LDH-B was introduced as a lens protein in these vertebrates.

τ-Crystallin is a major component of the lenses of lampreys, some fish, reptiles and birds (8, 11). In duck, a single-copy gene gives rise to this abundant lens protein, and sequence analysis revealed that τ-crystallin is in fact α-enolase, a highly conserved, important glycolytic enzyme (45, 51). As a consequence, lenses containing τ-crystallin display considerable enolase activity, but the specific activity is greatly reduced as compared to the non-lens enzyme, probably because of age-related posttranslational modifications (51). Apparently, the enzyme has not been recruited for its catalytic activity but for some distinct structural property. Noteworthy in this respect is the finding that in yeast an enolase is found that also acts as a heat-shock protein (52).

In frog lenses a major lens protein, called ρ-crystallin, has been found (9) that turned out to be a member of the same protein superfamily as aldose reductase, aldehyde reductase (53), and prostaglandin F synthase (54). As yet, however, no enzymatic activity or non-lenticular occurrence has been reported for ρ-crystallin.

Also in the mammalian lenses examples of evolutionary tinkering (19) are found (13, 44). In rabbit and hare lenses a 35 kDa protein, named λ-crys-

tallin, has been identified (13). Sequence data revealed that this λ-crystallin is distantly related to hydroxyacyl CoA dehydrogenases. It is not only found in lenses, where it accounts for 8% of the total protein content, but low expression levels can be detected in other tissues (13). Recently, a unique 36 kDa protein was isolated from guinea pig lenses. It was called ζ-crystallin, and turned out to be homologous to alcohol dehydrogenase (44). Obviously, the recruitment of enzymes or related proteins as lens structural proteins is an evolutionary strategy which has been applied independently in different lineages.

Outline of this Study

Although a wide variety of experiments on lenses and lens proteins have provided some clues about lens transparency (5, 6, 55, 56) the challenge remains to unravel the physicochemical properties that make crystallins suited for functioning in the lens. To approach this structure-function relationship, the complete structure of the hamster αA-crystallin gene was determined (chapter 2), and the functional constraints that shaped its product were investigated (chapter 3). In some mammals an additional α-crystallin subunit, which differs from the other α-chains due to an insertion of 23 residues in the middle of the subunit, is present (chapter 4). The effect of this insert peptide on the integrity of the α-crystallin aggregate, and on the spatial conformation of the subunit was examined (chapter 5). Finally, in chapters 6 and 7, some crystallins were identified as being identical with or related to certain enzymes. In the case of ε-crystallin/LDH-B4 the enzyme was investigated for possible adaptations to the additional function in the lens (chapter 6). For the time being, the structure-function relationships of the crystallins remain unclear. However, the important discovery that some crystallins are in fact active enzymes, has provided the possibility to investigate lens protein parameters like stability and turnover by monitoring their enzymatic activity.

REFERENCES

1 Bloemendal, H. (1981) Molecular and Cellular Biology of the Eye Lens (Wiley, New York).

2 Piatigorsky, J. (1981) Differentiation 19, 134-153.

3 Harding, J.J. & Crabbe, M.J.C. (1984) in The Eye, ed. Davson, H. (Acad. Press, London, New York) 3rd edn., pp. 207-492.

4 Maisel, H. (1985) The Ocular Lens (Marcel Dekker, New York).

5 Slingsby, C. (1985) Trends Biol. Sci. 10, 281-284.

6 Tardieu, A. & Delaye, M. (1988) Ann. Rev. Biophys. Biophys. Chem. 17, 47-70.

7 Piatigorsky, J. (1984) Cell 38, 620-621.

8 Stapel, S.O. & de Jong, W.W. (1983) FEBS Lett. 162, 305-309.

9 Tomarev, S.I., Zinovieva, R.D., Dolgilevich, S.M., Luchin, S.V., Krayer, A.S., Skryabin, K.G. & Gausse, G.G. jr (1984) FEBS Lett. 171, 297-302.

10 Stapel, S.O., Zweers, A., Dodemont, H.J., Kan, J.H. & de Jong, W.W. (1985) Eur. J. Biochem. 147, 129-136.

11 Williams, L.A., Ding, L., Horwitz, J. & Piatigorsky, J. (1985) Exp. Eye Res. 40, 741-749.

12 Huang, Q.L., Russell, P., Stone, S.H. & Zigler, J.R. jr (1987) Curr. Eye Res. 6, 725-732.

13 Mulders, J.W.M., Hendriks, W., Blankesteijn, W.M., Bloemendal, H. & de Jong, W.W. (1988) J. Biol. Chem. 263, 15462-15466.

14 de Jong, W.W. & Hendriks, W. (1986) J. Mol. Evol. 24, 121-129.

15 de Jong, W.W. (1981) in Molecular and Cellular Biology of the Eye Lens, ed. Bloemendal, H. (Wiley, New York) pp. 221-278.

16 Wilson, A.C., Carlsson, S.S. & White, T.J. (1977) Annu. Rev. Biochem. 46, 573-639.

17 Kimura, M. (1983) The Neutral Theory of Molecular Evolution (Cambridge Univ. Press, Cambridge, England).

18 Nei, M. & Koehn, R.K. (1983) Evolution of Genes and Proteins (Sinauer, Sunderland, MA).

19 Jacob, F. (1983) in Evolution from Molecules to Men, ed. Bendall, D.S. (Cambridge Univ. Press, Cambridge, England) pp. 131-144.

20 Wistow, G. & Piatigorsky, J. (1988) Annu. Rev. Biochem. 57, 479-504.

21 van den Oetelaar, P.J.M., Clauwaert, J., van Laethem, M. & Hoenders, H.J. (1985) J. Biol. Chem. 260, 14030-14034.

22 Thomson, J.A. & Augusteyn, R.C. (1988) Curr. Eye Res. 7, 563-569.

23 de Jong, W.W., Mulders, J.W.M., Voorter, C.E.M., Berbers, G.A.M., Hoekman, W.A. & Bloemendal, H. (1988) in Advances in Posttranslational Modifications of Proteins and Aging, eds. Zappia, V., Galletti, P., Porta, R. & Wold, F. (Plenum, New York) Adv. Exp. Med. Biol. 231, 95-108.

24 Hendriks, W., Sanders, J., de Leij, L., Ramaekers, F., Bloemendal, H. & de Jong, W.W. (1988) Eur. J. Biochem. 174, 133-137.

25 Li, L.K. & Spector, A. (1974) Exp. Eye Res. 19, 49-57.

26 Siezen, R.J. (1981) FEBS Lett. 133, 1-8.

27 Siezen, R.J. & Argos, P. (1983) Biochim. Biophys. Acta 748, 56-67.

28 Wistow, G. (1985) FEBS Lett. 181, 1-6.

29 Thomson, J.A. (1985) α-Crystallin: a Model for Multisubunit Protein Assemblies, PhD Thesis, University of Melbourne, Australia.

30 Augusteyn, R.C. & Koretz, J.F. (1987) FEBS Lett. 222, 1-5.

31 Bindels, J.G., Siezen, R.J. & Hoenders, H.J. (1979) Ophthalmic Res. 11, 441-452.

32 Tardieu, A., Laporte, D., Licinio, P., Krop, B. & Delaye, M. (1986) J. Mol. Biol. 192, 711-724.

33 Stapel, S.O., Leunissen, J.A.M., Versteeg, M., Wattel, J. & de Jong, W.W. (1984) Nature 311, 257-259.

34 de Jong, W.W., Leunissen, J.A.M., Leenen, P.J.M., Zweers, A. & Versteeg, M. (1988) J. Biol. Chem. 263, 5141-5149.

35 Ingolia, T.D. & Craig, E.A. (1982) Proc. Natl. Acad. Sci. USA 79, 2360-2364.

36 Nene, V., Dunne, D.W., Johnson, K.S., Taylor, D.W. & Cordingley, J.S. (1986) Mol. Biochem. Parasitol. 21, 179-188.

37 Schuster, G., Even, D., Kloppstech, K. & Ohad, I. (1988) EMBO J. 7, 1-6.

38 de Jong, W.W., Leunissen, J.A.M., Hendriks, W. & Bloemendal, H. (1988) in Molecular Biology of the Eye: Genes, Vision & Ocular Disease, eds. Piatigorsky, J., Zelenka, P. & Shinohara, T. (Alan Liss, New York) UCLA Symp. Mol. Cell. Biol. 88, in press.

39 Driessen, H.P.C., Herbrink, P., Bloemendal, H. & de Jong, W.W. (1981) Eur. J. Biochem. 121, 83-91.

40 van Rens, G., Raats, J., Driessen, H.P.C., Meera Khan, P., de Jong, W.W.

& Bloemendal, H. (1988) Gene, in press.

41 Wistow, G., Summers, L. & Blundell, T. (1985) Nature 316, 771-773.

42 Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. & Wistow, G. (1981) Nature 289, 771-777.

43 Summers, L., Slingsby, C., White, H., Narebor, M., Moss, D., Miller, L., Mahadevan, D., Lindley, P., Driessen, H. & Blundell, T. (1984) in Human Cataract Formation, eds. Nugent, J. & Whelan, J. (Pitman, London) CIBA Foundation Symposium 106, 219-236.

44 Borras, T. & Rodokanaki, A. (1988) Proc. Int. Soc. Eye Res. 5, 39.

45 Wistow, G. & Piatigorsky, J. (1987) Science 236, 1554-1556.

46 Yeh, L.-S.L., Elzanowski, A., Hunt, L.T. & Barker, W.C. (1988) Comp. Biochem. Physiol. 89B, 433-437.

47 Piatigorsky, J., O'Brien, W.E., Norman, B.L., Kalumuck, K., Wistow, G.J., Borras, T., Nickerson, J.M. & Wawrousek, E.F. (1988) Proc. Natl. Acad. Sci. USA 85, 3479-3483.

48 Wistow, G., Mulders, J.W.M. & de Jong, W.W. (1987) Nature 326, 622-624.

49 Hendriks, W., Mulders, J.W.M., Bibby, M.A., Slingsby, C., Bloemendal, H. & de Jong, W.W. (1988) Proc. Natl. Acad. Sci. USA 85, 7114-7118.

50 Wilson, A.C., Kaplan, N.O., Levine, L., Pesce, A., Reichlin, M. & Allison, W.S. (1964) Fed. Proc. Fed. Am. Soc. Exp. Biol. 23, 1258-1266.

51 Wistow, G., Lietman, T., Williams, L., Stapel, S.O., de Jong, W.W., Horwitz, J. & Piatigorsky, J. (1988) J. Cell Biol., in press.

52 Iida, H. & Yahara, I. (1985) Nature 315, 688-690.

53 Carper, D., Nishimura, C., Shinohara, T., Dietzchold, B., Wistow, G., Craft, C., Kador, P. & Kinoshita, J.H. (1987) FEBS Lett. 220, 209-213.

54 Watanabe, K., Fujii, Y., Nakayama, K., Ohkubo, H., Kuramitsu, S., Kagamiyama, H., Nakanishi, S. & Hayaishi, O. (1988) Proc. Natl. Acad. Sci. USA 85, 11-15.

55 Benedek, G.B. (1984) in Human Cataract Formation, eds. Nugent, J. & Whelan, J. (Pitman, London) CIBA Foundation Symposium 106, 237-247.

56 Siezen, R.J., Fisch, M.R., Slingsby, C. & Benedek, G.B. (1985) Proc. Natl. Acad. Sci. USA 82, 1701-1705.

CHAPTER 2

COMPLETE STRUCTURE OF THE HAMSTER αA-CRYSTALLIN GENE: REFLECTION OF AN EVOLUTIONARY HISTORY BY MEANS OF EXON SHUFFLING

Richard van den Heuvel, Wiljan Hendriks,

Wim Quax & Hans Bloemendal

J. Mol. Biol. 185 (1985) 273-284.

Complete Structure of the Hamster αA Crystallin Gene

Reflection of an Evolutionary History by Means of Exon Shuffling

Richard van den Heuvel, Wiljan Hendriks, Wim Quax and Hans Bloemendal†

Department of Biochemistry, University of Nijmegen
Geert Groote Plein Noord 21
6525 EZ Nijmegen, The Netherlands

(Received 28 January 1985, and in revised form 8 May 1985)

The eye lens contains a structural protein α crystallin composed of two homologous primary gene products αA_2 and αB_2 In certain rodents still another α crystallin polypeptide αA^{Ins} occurs which is identical to αA_2 except that it contains an insertion peptide between residues 63 and 64 In this paper we describe the complete αA crystallin gene that has been cloned from DNA isolated from Syrian golden hamster Evidence is provided that the αA gene is present as a single copy in the hamster genome The detailed organization of the gene has been established by means of DNA sequence analysis and S_1 nuclease mapping revealing that the gene consists of four exons The first exon contains the information for the 68 base-pair long 5′ non coding region as well as the coding information for the first 63 amino acids The second exon encodes the 23 amino acid insertion sequence, the third exon codes for amino acid 87 to 127 of the αA^{Ins} chain, whereas the last exon encodes the C terminal 69 amino acids and contains the information for the 523 base pair long 3′ non coding region The second exon is bordered by a 3′ splice junction (A G/G C), which deviates from the consensus for donor splice sites (A G/G T) This deviation is found in both hamster and mouse An internal duplication was detected in the first exon by using a DIAGON generated matrix for comparison By means of similar DIAGON-generated matrices it was confirmed that the amino acids coded for by the third and fourth exons are homologous to the small heat shock proteins of *Drosophila Caenorhabditis* and soyabean The implications of the differential splicing and the evolutionary aspects of the detected homologies are discussed

1. Introduction

The structural proteins of the vertebrate eye lens, the crystallins, have been divided into four main classes that are immunologically distinct (for reviews, see Harding & Dilley 1976, Bloemendal 1977) α Crystallin is the predominant biosynthetic product in the adult mammalian lens and represents an example of controlled gene expression in terminally differentiated cells (Bloemendal, 1981) This high molecular weight protein is composed of two types of primary gene products, αA_2 and αB_2, which show about 60% sequence homology, suggesting that an ancient duplication of an ancestral gene gave rise to separate αA and αB genes (van der Ouderaa *et al*, 1974) cDNAs have been isolated for rat, mouse, frog and chicken αA_2 crystallin (Dodemont *et al*, 1981 King *et al* 1982, Tomarev *et al* 1983, Yasuda *et al*, 1982) To date, only in the rodent families *Muridae* (rat mouse) and *Cricetidae* (hamster, gerbil) has an additional αA crystallin subunit been detected, i e αA^{Ins}, which is identical to αA_2 except for a stretch of 22 amino acids inserted between residues 63 and 64 of the αA_2 chain (Cohen *et al*, 1978, de Jong *et al*, 1980) It was suggested that the αA_2 and αA^{Ins} mRNAs result from differential splicing of the same primary transcript (Bloemendal *et al*, 1978, de Jong *et al*, 1980) This idea was supported by strong indications for the existence of just one αA crystallin gene in the mouse This gene gives rise to both the αA_2 and the αA^{Ins} mRNAs which are formed in a ratio of approximately 10 1 in the mouse as well as in the hamster (King & Piatigorsky, 1984, Dodemont 1984) Furthermore, studies on the structure of a partial genomic mouse αA clone revealed that the insertion in αA^{Ins} is

† Author to whom requests for reprints should be sent

coded for by a sequence of 69 nucleotides inside the first intron This sequence is bordered by a 3′ splice junction (A·G G·C), which deviates from the consensus for donor splice sites (A·G/G·T). The deviation has been proposed as a possible explanation for the differential splicing of the αA primary transcript (King & Piatigorsky, 1983)

Another very interesting observation concerning the α crystallins is the homology that was found between the calf αB_2 chain and the small heat-shock proteins of *Drosophila* (Ingolia & Craig, 1982) Recently, homology has also been found between small heat shock proteins of *Caenorhabditis* and soyabean on the one hand, and αB_2 crystallin on the other hand (Russnak *et al*, 1983, Schöffl *et al*, 1984) We investigated the problem of whether there is a correlation between the αA gene structure and the region of sequence similarity

2. Materials and Methods

(a) *Construction and screening of a hamster genomic library*

High molecular weight DNA isolated from the liver of Syrian golden hamster (*Mesocricetus auratus*) was digested with *Sau*3A and the fraction ranging from 15 to 20 ($\times 10^3$) base-pairs was cloned in Charon 28 (Rimm *et al*, 1980) We screened the obtained library using previously isolated αA_2-specific cDNAs from rat and calf (Dodemont *et al* 1981 Quax Jeuken *et al*, 1985) Hybridization was carried out at 42°C during 16 h in 35% formamide 5×SSC (SSC is 0·15 M-NaCl, 0·015 M-sodium citrate pH 7·0) 1×Denhardt's solution, 20 mM-sodium phosphate (pH 6 8) 5 mM-EDTA, 100 μg ml^{-1} single stranded salmon sperm DNA, and 30 μg ml^{-1} single stranded *Escherichia coli* DNA Washings were done at 42°C with the hybridization mixture and at 50°C for 30 min in 2×SSC, 0 1% sodium dodecylsulphate Positive plaques were purified, and characterization was done by restriction enzyme mapping using established methods

(b) *DNA sequence analysis*

A 5 6 kb† *Eco*RI-*Sst*I fragment, derived from clone λHaαACr3, containing that part that hybridizes to our cDNA probes, was subcloned into a suitable vector This vector is derived from plasmid pBR322 and contains an *Eco*RI and a *Sst*I site (a gift from Dr P van Weezenbeek) The 5·6 kb *Eco*RI-*Sst*I insert of this subclone was digested with different enzymes (*Sau*3A *Hae*III and *Hpa*II) and the resulting fragments were cloned into M13mp8 mp9 mp10 or mp11 (Messing 1983) For some stretches, known restriction fragments were directly subcloned The chain termination method (Sanger *et al*, 1980) was used to sequence the M13 inserts Analysis of the sequence reaction products was done on 40 cm length gels of 6% (w/v) polyacrylamide/7 M-urea

(c) *Computer programs*

The sequence gel readings were recorded, compared, edited and assembled by the 'shotgun' programs of Staden (1977) The sequence comparisons were done by the DIAGON program (Staden 1982) which is based on the procedure of McLachlan (1971) and run on a VAX 11/780 computer The scoring system for protein comparison is based on the frequency of substitution of one amino acid by another in a number of protein families (Schwartz & Dayhoff 1978) Only scores above a preset threshold appear on the display as a dot corresponding to the midpoint of the compared blocks The blocklength and thresholds used as well as the corresponding probabilities are given in the Figure legends and the text

(d) *S_1 nuclease analysis*

Total RNA was extracted from hamster lenses by homogenization in 6 M guanidinium isothiocyanate 10% Sarcosyl The solution was then centrifuged through a 5·7 M-CsCl cushion, the RNA pellet was dissolved and precipitated with ethanol The 5′ and 3′ ends of the gene were subcloned into M13mp8 and a new ^{32}P-labelled strand was synthesized from the universal primer The newly synthesized strand was hybridized to 1 μg of total hamster lens RNA for 18 h at 55 C in 80% formamide, 400 mM-NaCl 40 mM-PIPES (pH 6·4) and 1 mM-EDTA The hybrids were directly diluted in 300 μl of S_1 nuclease solution containing 300 mM sodium acetate (pH 4 5), 0 12 mM-$ZnSO_4$ 100 mM-NaCl and 300 units of S_1 nuclease (modification of Berk & Sharp, 1977) S_1 nuclease digestion was performed for 45 min at 30°C The S_1-resistant material was directly precipitated with ethanol and analysed on a 6% polyacrylamide sequence gel Sequence reactions of known fragments were used as size markers

3. Results

(a) *Cloning and identification of the complete hamster αA crystallin gene*

High molecular weight DNA from Syrian golden hamster liver was partially digested with *Sau*3A and cloned into the λ vector Charon 28. After *in vitro* packaging and transformation about 500,000 plaques were obtained. This hamster genomic library was screened using an αA_2-specific cDNA with a length of approximately 1000 base-pairs isolated from a bovine cDNA library (Quax-Jeuken *et al*, 1985) and with a rat αA_2 cDNA clone (Dodemont *et al*, 1981) Three Charon clones containing αA homologous sequences were obtained (Fig 1) The fragments resulting from an *Eco*RI/*Sst*I digestion of the Charon phage λHaαACr3 were cloned into a suitable vector derived from pBR322. One of the resulting subclones, pHCr3 (Fig 1) was used for further analysis, since all sequences hybridizing to our αA_2 cDNA clones were located within pHCr3 To ensure that all genomic αA sequences were obtained, both genomic Syrian golden hamster liver DNA and λHaαACr3 were digested with *Kpn*I and *Eco*RI/*Kpn*I, followed by gel electrophoresis and blotting Upon screening with the bovine αA_2 cDNA probe, the *Kpn*I digest shows one hybridizing band of 7·1 kb in both DNAs, and the *Eco*RI/*Kpn*I double digest shows one hybridizing band of 5·4 kb in both DNAs (Fig 2) This provides evidence that all

† Abbreviations used kb 10^3 base-pairs, bp, base-pairs

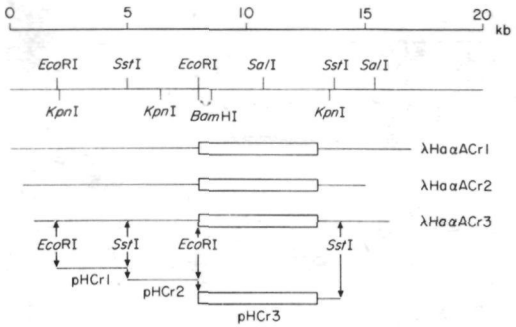


Figure 1. Cloning and subcloning of the hamster αA crystallin gene. The αA crystallin gene of the Syrian golden hamster was cloned into the λ-derived vector Charon 28. The resulting clones λHaαACr1 to 3 were subjected to restriction analysis. The physical map of the αA gene and flanking sequences in hamster genomic DNA is given in the second line. The open boxes represent those regions in which we found hybridization to our αA cDNAs. The fragments resulting from an *Eco*RI/*Sst*I digestion of the Charon phage λHaαACr3 were subcloned into a suitable vector derived from pBR322. The resulting subclones were named pHCr1 to 3.

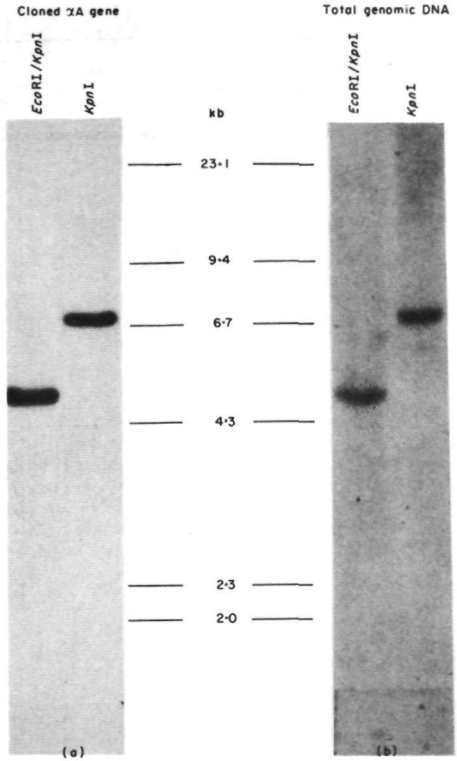


Figure 2. Comparison of λHaαCr3 with total hamster DNA. (a) λHaαACr3 and (b) total hamster DNA were digested and double-digested with the indicated enzymes. *Hin*dIII digested λ DNA was used as marker. Southern blots were hybridized with the ^{32}P-labelled rat cDNA. Hybridization was carried out at 42°C during 16 h in 50% formamide, 5 × SSC, 3 × Denhardt's solution, 20 mM-sodium phosphate (pH 6·8), 5 mM-EDTA and 100 μg ml^{-1} single-stranded salmon sperm DNA. Washings were done at 42°C with the hybridization mixture, at 42°C for 15 min in 2 × SSC, 0·5% sodium dodecyl sulphate (SDS) and lastly at 55°C for 30 min in again 2 × SSC, 0·5% SDS.

genomic αA sequences are situated within λHaαACr3, and since all sequences hybridizing to our αA cDNAs are located in pHCR3, all genomic αA sequences must be present in pHCR3. Detailed analysis of pHCr3 (see below) has revealed that it contains only one gene. Therefore, we conclude that only one αA gene is present in the genome of the hamster.

(b) *Determination of the complete structure of the αA gene*

In order to elucidate the structure of the hamster αA gene the nucleotide sequence of the major part of pHCr3 was established by means of the dideoxy sequencing method (Sanger *et al.*, 1980; Messing *et al.*, 1981). Accurate nucleotide sequences were obtained by sequencing both DNA strands at least once. Computer analysis of all determined sequences revealed the existence of four exons. We numbered these exons 1 to 4, which is a slightly different nomenclature from that used by King & Piatigorsky (1983). We propose to call the optionally spliced sequence coding for the insertion an exon, as is common for differentially spliced sequences. The structure of the hamster αA crystallin gene is presented in Figure 3. The determined sequence is depicted in the upper line of Figure 5. The first exon contains the coding information for the first 63 amino acids. The insertion sequence found in approximately 10% of the hamster and mouse αA_2 mRNAs (Dodemont, 1984; King & Piatigorsky, 1984) was found to be encoded by exon 2. This second exon has a length of 69 bp and encodes 23 amino acids. The amino acid sequence derived from this exon differs in three positions from the amino acid sequence of the hamster αA^{Ins} chain previously proposed (de Jong *et al.*, 1980). A tryptophan residue was now found to be present between methionine at position 5 and the following phenylalanine of the insertion. Furthermore, methionine and histidine at positions 9 and 10, respectively, were previously reported to be inverted. Compared to the mouse, one different amino acid was found; at the position of the last isoleucine of the hamster insertion sequence a valine residue was found in the mouse. It is remarkable that in the same position in the rat insertion sequence yet another amino acid, glycine,

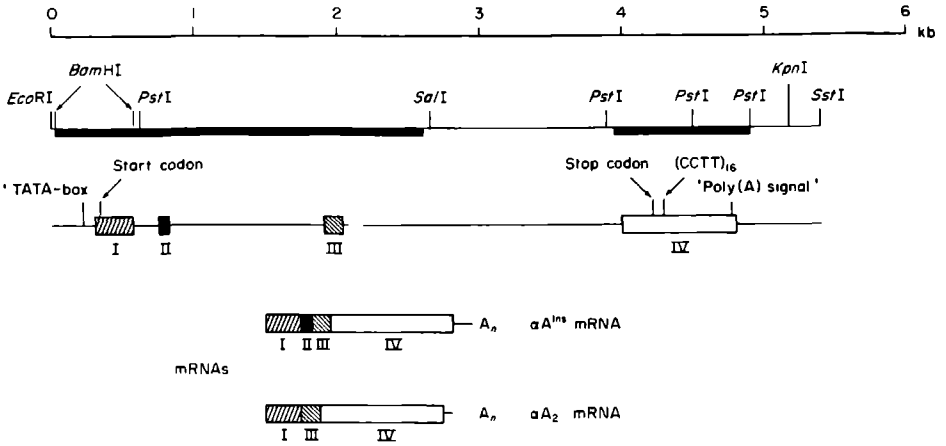


Figure 3. Physical map of the hamster αA crystallin gene The second line depicts the sites for some restriction enzymes that were often used The black boxes indicate those regions for which the nucleotide sequence was established On the third line the sizes and positions of the αA exons are represented by boxes and numbered with Roman numerals The position of transcription and translation signals is given as well as the position of a remarkable 16 times C C T T repeat The last two lines represent the two mRNAs that are transcribed from the αA gene

is found (Cohen *et al* 1978) The difference between hamster and mouse sequences is caused by a single substitution in the first position of the corresponding codon Another single substitution in the second position of the same codon can explain the difference between mouse and rat αA^{Ins} The reason for the apparent high substitution rate for this particular residue is unclear

The third exon has a length of 123 bp and contains the coding information for amino acids 87 up to 127 of the αA^{Ins} chain The fourth exon is 730 bp in length and contains the coding information for the last 69 amino acids as well as the long 3 non coding region

(c) *Precise mapping of the transcribed region*

The exact length of the first exon was established by means of S_1 nuclease mapping The 650 bp pHCr3 fragment from the *Eco*RI to the second *Bam*HI restriction site containing the 5 end of the gene was cloned into M13mp8 The template was isolated and ^{32}P labelled transcripts using the universal primer were synthesized The newly synthesized strand was hybridized to total hamster lens RNA and after S_1 nuclease digestion the sample was run on a 6% polyacrylamide sequence gel (Fig 4(a)) In a similar way the localization of the poly(A) addition site was determined This time an M13mp8 clone containing the 3' end of the gene which is located within a pHCr3 fragment ranging from the third to the fourth *Pst*I site was transcribed (from the fourth to the third *Pst*I site) and the synthesized strand was hybridized to total hamster lens RNA treated with S_1 nuclease and analysed (Fig 4(b))

In this way it was established that the first exon has a length of 257 bp, leaving a length of 68 bp for the 5 non coding region The first exon is preceded by a T A T A T A A sequence which is homologous to the consensus promoter for RNA transcription in eukaryotic genes (Goldberg, 1979) However, no consensus ' CAAT box was found The transcription start position (CAP site), is found at position 350 in Figure 5 and confirms the rule that the CAP site is preferably A preceded by C, and mostly situated around 31 bp downstream from the TATA box

The length of the 3' non coding region was found to be 523 bp The consensus poly(A) addition signal A A T A A A is located 20 bp upstream from the poly(A) addition site Furthermore a remarkable 16 times C C T T repeat was detected within the 3' non coding region, which is absent in the mouse cDNA (Fig 5(b)) Since it has been shown previously that after removal of the poly(A) track the αA_2 mRNAs of hamster and mouse have an equal length of 1110 ± 20 bp (Dodemont, 1984) the repeat is most probably present in the hamster αA mRNAs Simple repetitive DNA sequences can be transcribed into mature mRNA in eukaryotes (Tautz & Renz 1984) and can even be translated into protein (Quax Jeuken *et al*, 1984) The interesting idea has been put forward that such sequences are often involved in recombinational events (Slightom *et al* 1980) Such an event might be the reason why this sequence is not found in the mouse αA_2 mRNA As another example of a simple sequence the stretch of 17 T residues beginning at position 1050 (Fig 5(a)), should be mentioned This sequence has also not been found in the mouse αA crystallin gene

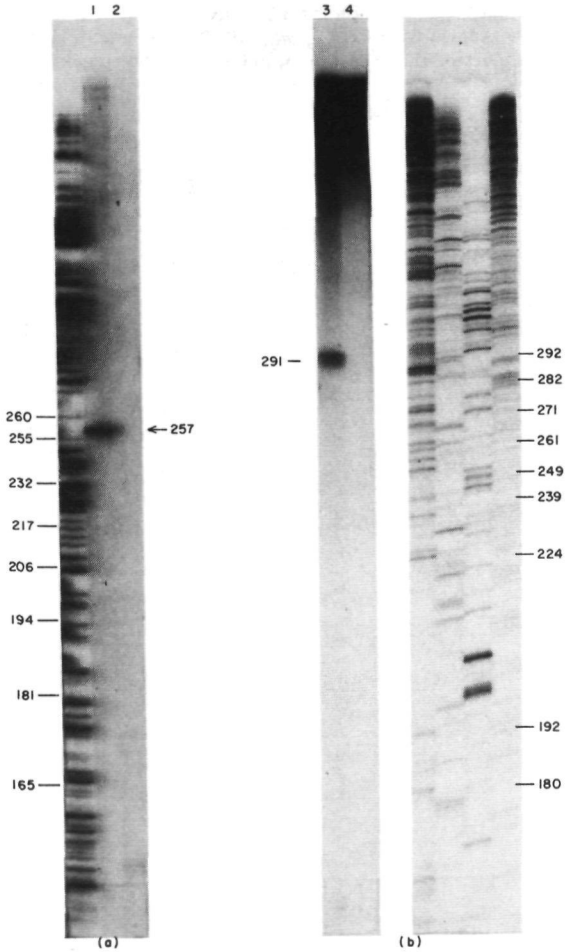


Figure 4. Determination of the exact size (in nucleotides) of the transcribed region. (a) Lane 1: an M13 subclone containing the 5′ part of the gene (described in the text) was transcribed from the universal primer. The ^{32}P-labelled newly synthesized strand was hybridized to total hamster lens RNA and treated with S_1 nuclease. The products were analysed on a polyacrylamide gel. A sequence reaction was used as a size marker. Lane 2: the same experiment without addition of RNA. (b) Lane 3: with the aid of an M13 subclone containing the 3′ part of the αA gene the position of the poly(A) addition signal was mapped (see the text). Lane 4: the same experiment without addition of RNA.

(d) *Splice junctions in the αA gene*

A comparison of the various splice junctions with the consensus splice sequences (Breathnach *et al.*, 1981; Mount, 1982) is given in Table 1. All splice junctions obey the consensus rules, except for the 3′ splice site of exon 2. In this case G and C were found to be the first two nucleotides in the intron instead of the highly conserved G and T. Exactly the same deviation was found in the mouse αA gene and it has been proposed that this deviation might lead to the differential splicing of the primary transcript (King & Piatigorsky, 1983, 1984). The fact that the same deviation has been found in hamster as well as in mouse strongly supports this idea. Whether this deviation is the sole reason for differential splicing to occur, or whether sequences within intron 2 play a role remains to be investigated. In this respect it is noteworthy that a similar A·G/G·C splice junction in a chicken α^D globin gene leads to normal splicing of the primary transcript (Dodgson & Engel, 1983).

(e) *Comparison of αA sequences of hamster and mouse*

In order to gather more information concerning the evolution of the αA crystallin gene, a

comparison between the nucleotide sequence of the αA gene of the hamster and the known 5′ part of the mouse αA gene plus the 3′ part of the mouse αA_2 cDNA clone (King & Piatigorsky, 1983, King *et al*, 1982) was made (Fig 5). Furthermore, we wondered whether regions that play a role in gene

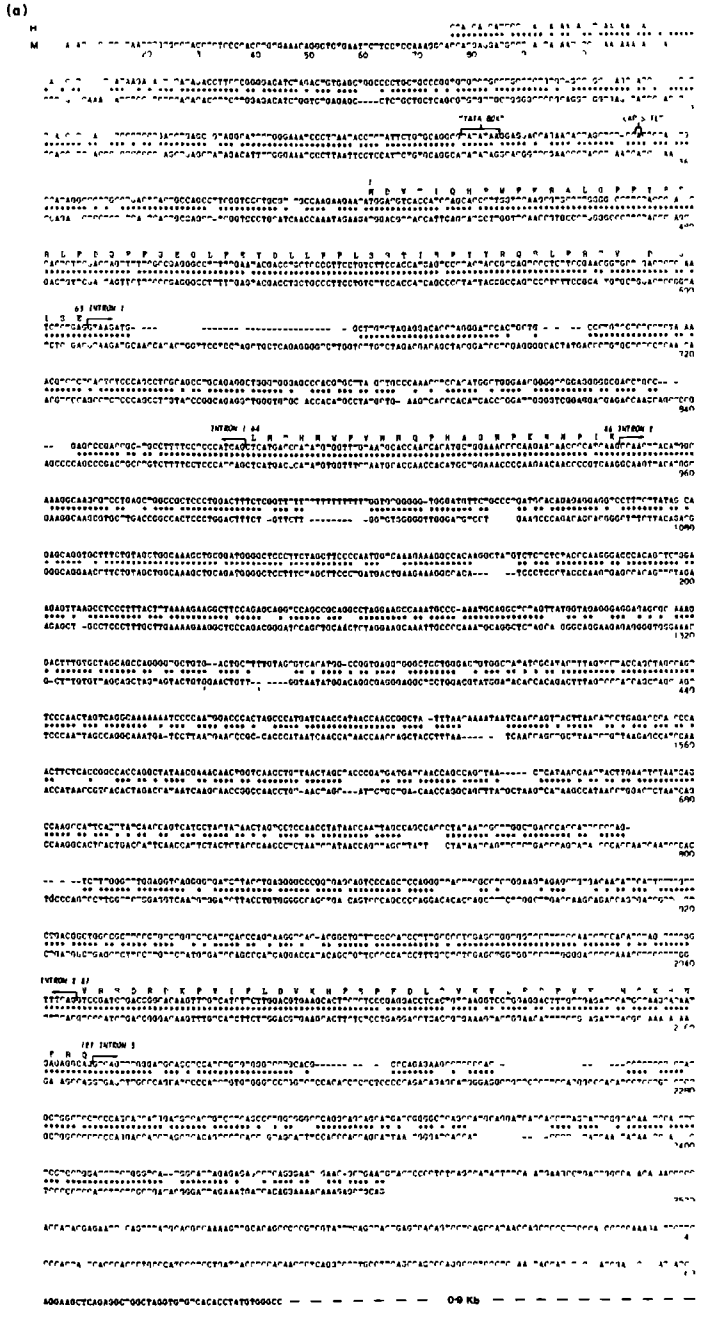


Fig. 5

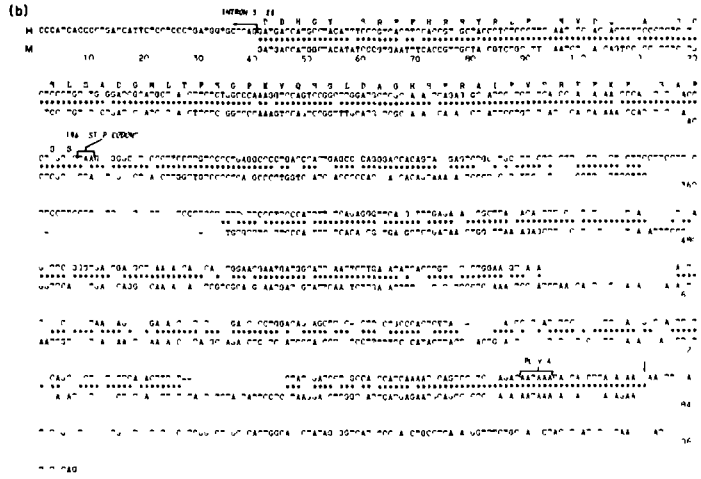


Figure 5. Sequence of the hamster (H) αA crystallin gene, compared with mouse (M) αA sequences. The sequence of the hamster αA crystallin gene is represented in the upper line in (a) and (b). (a) Represents the 5′ part of the gene in which the known part of the mouse gene is given in the lower line. Assumed deletions are given in both sequences by broken lines. Identical nucleotides in equal positions are indicated by asterisks. (b) Represents the 3′ part of the gene here the 3′ part of the mouse αA cDNA is given in the lower line. The amino acid sequence is given in single letter code above their corresponding codons. The intron positions are marked with brackets. The length of the part of intron 3 that was not sequenced is given in brackets. TATA BOX: region that is homologous to the Goldberg–Hogness box. CAP SITE: location of the capped nucleotide. POLY A: the location of the poly(A) addition signal. The arrow indicates the position of the poly(A) addition site. Sequence hyphens are omitted for clarity.

expression and its regulation could be identified. Such regions might be better-conserved than adjacent sequences. The alignment presented in Figure 5 was obtained by maximizing the homology and minimizing the number of single base pair deletions. If we turn our attention first to those regions for which there is clear evolutionary pressure, we indeed see extreme conservation between hamster and mouse. The coding sequences, omitting the silent positions, are perfectly conserved except for one codon within the insertion sequence. It is remarkable that this insertion, which is present in only a fraction of the αA chains, is conserved to such a high extent and also at the DNA level (Fig. 5(a)). This must mean that there is stringent evolutionary pressure on the insertion sequence. Also the sequences belonging to the splice junctions are conserved to a high degree (Table 1).

Recently, the idea has been put forward that the sequences downstream from the poly(A) addition signal are also under evolutionary pressure because of their role in the formation of mature mRNA (McDevitt *et al.*, 1984; Gil & Proudfoot, 1984). From Figure 5(b) it appears that the sequences downstream from the poly(A) addition signal are conserved between hamster and mouse, which is in accordance with this idea. For the remaining sequences, which can be considered to evolve freely, and which we would like to denote as selectively neutral sequences, we calculated the homologies as described in the legend to Table 2. We can also calculate the expected degree of homology for these sequences by using the formula $r = K^c/2T$, where $r = 5.37 \times 10^{-9}$ substitutions per site per year and r is taken to be constant for a long period of divergence (approx. 75×10^6 years) and for all eukaryotic genes at selectively neutral positions (Miyata *et al.*, 1982). In other studies the same or even higher values for r have been mentioned (Perler *et al.*, 1980; Jeffreys, 1982). T is the time of divergence between hamster and mouse which was assumed to be 35×10^6 years a minimum estimate based on paleontological evidence (Hartenberger, 1985). Using the formula and values mentioned above we expected to find a corrected homology of at most 62%. However, the lowest corrected homology found was 73.2%. A possible explanation for the high homologies that were found might be that within the compared areas there are still some regions that fulfil specific functions. The high homology between the silent positions (87%), for example, might indicate that codon usage puts rather high evolutionary pressure on these positions. Previous results with the globin genes also suggest conservation of these positions to a higher degree than expected (Miyata & Hayashida, 1981). In this connection it should be noted that within the crystallin families preferential codon usage has also been found (Quax-Jeuken *et al.*, 1984). Therefore, it is reasonable to assume that not all positions compared as being neutral are indeed free to evolve, but that certain regions are under

Table 1

Exon/intron junction of the αA crystallin gene

		Donor site		Acceptor site	
Consensus		(C A)G gt agt A g		ccccccccc c n ag G ttttttttttt t	
		Glu		*Leu*	
	H	C AG gtaaga		ttttcctcccatcag CTC	
EXON I			intron 1		EXON II
	M	G AC gtaaga		ttttcctcccatcag CTC	
		Glu		*Leu*	
		Lys		*Val*	
	H	A AG gcaagt		aggttttgggtttcag GTC	
EXON II			intron 2		EXON III
	M	A AG gcaagt		gggtttgg tttcag GTC	
		Lys		*Val*	
		Gln		*Asp*	
	H	C AG gtgagt		cctgatggtgcccag GAT	
EXON III			intron 3		EXON IV
	M	C AC gtgagt			
		Gln			

Nucleotides around the splice sites are compared with consensus donor (left) and acceptor (right) sequences (Mount 1982) Upper and lower lines represent hamster (H) and mouse (M) sequences respectively Exon sequences are given in capital letters Amino acids bordering the introns are given above the corresponding codons The ag at the 5 and the gt at the 3 end of an intron are usually present except in the donor site bordering exon II where a gc was found at the 3 end The other nucleotides do differ from those in the consensus sequence more often Hyphens are omitted from sequences for clarity

evolutionary constraint because of their special functions

(f) *Internal duplication within the first exon*

Since internal duplications in the β–γ crystallin family have been reported previously (Driessen *et al* 1981 Blundell *et al* 1981) we searched for such duplications within the hamster αA^{Ins} polypeptide chain using a DIAGON generated matrix and comparing the αA^{Ins} chain with itself (Fig 6) A clear internal duplication within the amino acids coded for by exon 1 is found Amino acids 3 to 32 are homologous with amino acids 33 to 62 This duplication is slightly different in position from the one proposed before for the bovine αA_2 chain (Barker *et al* 1978) We also searched for the duplication at the DNA level A DIAGON matrix was generated whereby exon 1 was compared with itself (data not shown) The nucleotide 432 to 496 region revelead a significant sequence similarity to the nucleotide 522 to 586 region and corresponds to the duplication found at the protein level Since the duplication found within the hamster αA^{Ins} sequence divides the coding region of the first exon into two equal parts we suggest that the first exon

Table 2

Homologies between the hamster αA crystallin gene and mouse αA sequences

Region	Length	Mismatch	K^o	K^c	Homology^o (%)	Homology^c (%)
5 Upstream	267	53	0 199	0 231	80 1	76 9
5 Non coding	66	12	0 182	0 208	81 8	79 1
Intron 1	181	36	0 199	0 231	80 1	76 9
Intron 2	984	219	0 223	0 264	77 7	73 6
Intron 3	231	52	0 225	0 268	77 5	73 2
3 Non coding	449	78	0 174	0 198	82 6	80 2
Silent sites	105	13	0 124	0 135	87 6	86 5

In the first column the regions that were compared are given The alignment given in Fig 5 was used in the calculations In the second column the length of the compared area is given In calculating the length all deletions were omitted as well as the sequences belonging to the splice junctions and the sequences downstream from the poly(A) addition signal The number of silent positions was derived by counting the number of amino acids that are encoded by a codon in which only the first 2 nucleotides determine the encoded amino acid and in which the third position can be A G C or T The homologies were calculated according to Miyata *et al* (1981) By dividing the number of mismatches by the length of the corresponding areas a K^o value was obtained K^c the K value corrected for multiple events was derived from K by using the formula $K^c = (-3/4)\ln[1-(4/3)K^o]$ From the corresponding K values the homologies were calculated by homology = $100-(100 \times K)$

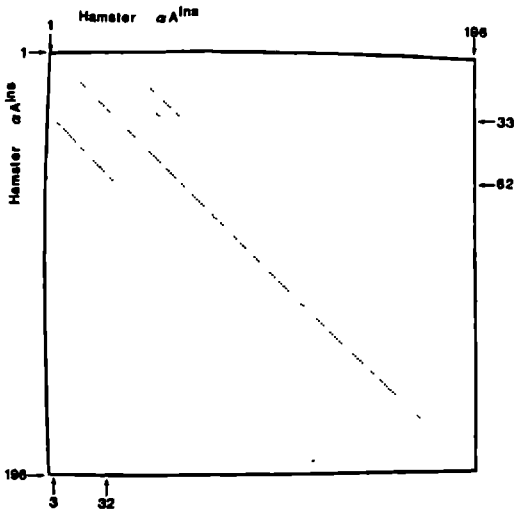


Figure 6. Internal duplication in the first exon The amino acid sequence of αA^{Ins} crystallin was compared with itself The comparison was done using the DIAGON program The blocklength was taken at 31 and the threshold was set at 331% This corresponds to a probability of 1.6×10^{-3} The arrows indicate the positions of the duplications, the numbers of the corresponding amino acids are indicated

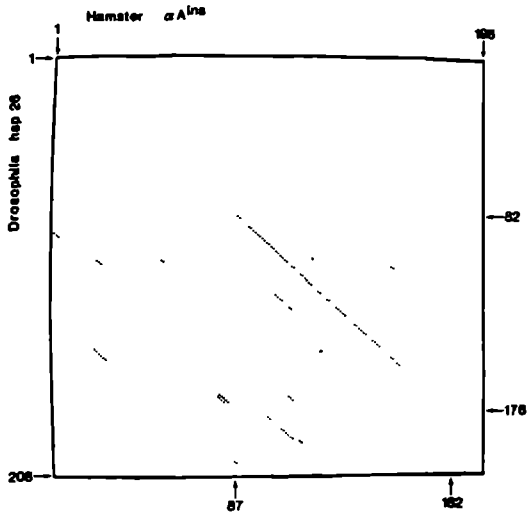


Figure 7. Homology between small heat shock proteins of *Drosophila* and αA^{Ins} crystallin The αA^{Ins} crystallin sequence was compared with the sequence given for *hsp*26 by Southgate *et al* (1983) The αA^{Ins} sequence is listed along the abscissa and the sequence of *hsp*26 along the ordinate The blocklength was taken at 19 and the threshold was set at 206% This corresponds to a probability of 3.7×10^{-3} The arrows mark the stretch of homology The numbers of the corresponding amino acids are given

has originated by duplication of an ancestral sequence encoding 30 amino acids The borders of the first exon in the present gene perhaps arose by sliding of the exon/intron junction Exon sliding has been found previously in another crystallin gene, namely the rat γ_{1-2} gene (Moormann *et al*, 1983)

(g) *The crystallin region homologous to the small* Drosophila *heat shock proteins maps at exon 3 and the coding region of exon 4*

In order to investigate, at the gene level, the nature of the homology found between the four small heat-shock proteins of *Drosophila*, two heat-shock proteins of soyabean and one heat-shock protein of the nematode *Caenorhabditis*, on the one hand, and bovine αB_2, on the other hand (Ingolia & Craig, 1982, Southgate *et al*, 1983, Schoffl *et al*, 1984, Russnak *et al*, 1983), we compared the hamster αA^{Ins} amino acid sequence with the amino acid sequence of *hsp*26 (Southgate *et al*, 1983) For this analysis the αA^{Ins} amino acid sequence was listed along the abscissa and the *hps*26 amino acid sequence along the ordinate The matrix generated is presented in Figure 7 A clear diagonal line was found representing sequence similarity between amino acids 87 to 182 of the αA^{Ins} chain, and amino acids 82 to 176 of the *hps*26 chain Similar results have been obtained with the other small heat shock proteins of *Drosophila* (data not shown) This region of the αA^{Ins} chain corresponds to the entire third exon and almost all of the coding region of the fourth exon It seems very unlikely that this sequence similarity has arisen by means of convergent evolution since evidence has been provided that preferential occurrence of structurally similar amino acids in analogous secondary structures does not give rise to a statistically significant sequence similarity for such a long stretch (Keim *et al*, 1981) Another argument against convergent evolution is that the homology with the small heat-shock proteins has now been shown for αB_2 as well as for αA_2 Since αA_2 and αB_2 are just 60% homologous, it seems very unlikely that the small heat-shock proteins converged to a sequence homologous to both αA_2 and αB_2 crystallin Therefore, the observed homology between heat shock proteins from a wide range of eukaryotes and α crystallin is most probably due to a common ancestral gene for these proteins

4. Discussion

(a) *The insertion peptide is encoded by an optional exon*

The vertebrate αA_2 chain has been subjected to extensive comparative protein sequence analysis

(see de Jong *et al* 1984) The obtained sequences have provided considerable information about the evolutionary history of αA, as well as suggestions concerning the origin of αA^{Ins} (de Jong *et al*, 1980) Studies on a mouse partial genomic αA clone and an αA^{Ins} specific cDNA of the mouse have supported the idea that αA_2 and αA^{Ins} arise from the same gene, by differential splicing of the primary transcript (King & Piatigorsky, 1983 1984) Our present finding that the G C flanking the second exon is present in the mouse as well as in the hamster αA gene provides new evidence for this idea The differential splicing of this gene elicits some ideas concerning the mechanism of splicing Recently, a lariat mechanism has been proposed, suggesting that a cut at the donor site is made simultaneously with or just before the formation of a 2′ to 5′ diester bond between the G at the splice junction and an A around 20 to 40 bp upstream from the acceptor site (Ruskin *et al* 1984, Padgett *et al*, 1984) Previous studies have led to the proposal that U1 RNA might be involved in splicing by bringing the donor and acceptor sites together (Rogers & Wall 1980 Lerner *et al*, 1980) Very recently, Tatei *et al* (1984) have shown that the U1 RNA–protein complex does indeed bind specifically to both 5′ and 3′ splice junction sequences in RNA If we try to reconcile the differential splicing of the αA primary transcript with this splicing model, we see that in this particular case intron 1 cannot be spliced before intron 2 If intron 1 were to be spliced before intron 2 it cannot be explained why exon 1 is not always spliced to exon 2, since exon 2 has a 5′ splice site that obeys the consensus perfectly and is functional although only in about 10% of the cases This suggests that the introns in the αA primary transcript are removed in a 3′ to 5′ order although a concerted mechanism cannot be ruled out At least it is clear that the splicing of intron 1 is affected by the splicing of intron 2 There are more examples where splicing in one intron affects splicing in another intron For instance Medford *et al* (1984) have shown recently that such an effect exists in the case of the splicing of troponin pre-mRNA In this case the seventh (α) and eighth (β) exons are differentially spliced in a mutually exclusive way

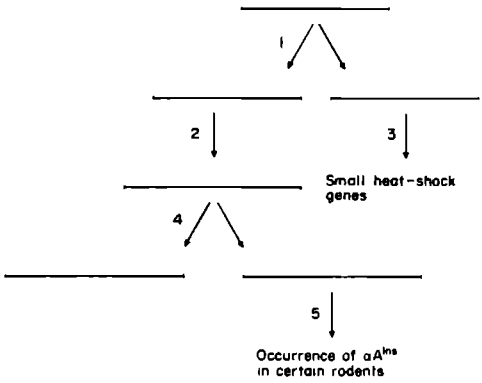


Figure 8. Tentative model for the evolution of the αA crystallin gene 1 Duplication of the most primitive ancestral gene, in which sequences homologous to the first exon of the present αA gene were most probably not present (see the text) that gave rise to an ancestral gene for the small heat shock proteins and an ancestral gene for the α crystallins This event must have taken place before the divergence of plants and animals because small heat shock proteins of both plants and animals are homologous to the α crystallins 2 Internal duplication within the first exon 3 Gene duplications that gave rise to the small heat shock protein gene family 4 Gene duplication that gave rise to separate αA and αB genes 5 Occurrence of an insertion sequence in certain rodents This event must have taken place before the divergence of *Muridae* and *Cricetidae*

Furthermore we conclude that in the process of selection of the correct donor and acceptor sites, the 3′ splice site of exon 2 is recognized as being a good donor site in only about 10% of the cases, resulting in the αA^{Ins} mRNA In the rest of the cases the donor site of exon 2 is not recognized and the donor site of exon 1 is used, giving rise to the αA_2 mRNA A possible explanation might be that base-pairing of the G C with U1 RNA is less efficient than base-pairing of the normal G T with U1 RNA The availability of the complete αA genomic clone will enable us to study the unique (so far) differential splicing mechanism of the αA pre mRNA in more detail

(b) *Is there a correlation between the exon/intron distribution in the αA crystallin gene and protein domains?*

It has been proposed that exons might play a role in the evolution of proteins and that therefore the exon/intron distribution of a gene might show a correlation with protein domains and protein structure (Gilbert, 1978, Blake, 1978) In contrast to the extensive information concerning the primary structure of αA little information has been obtained so far on the secondary and tertiary structures However, recently a six-folding unit model was proposed for the bovine αA_2 chain (Siezen & Argos 1983) Using this model we found that the end of exon 1 maps exactly with the end of folding unit 2 and the end of exon 3 maps almost at the end of folding unit 4 This observation favours the assumption that each exon encodes two folding units, a situation that is also found in the γ crystallins (Moormann *et al*, 1983) The second exon encodes a separate unit the insertion sequence As for the position of the insertion it is noteworthy that from a recent examination of intron positions in the genes of serine proteinases and dihydrofolate reductases it was found that intron positions frequently occur at surface regions where the length of the polypeptide chains is significantly altered by addition or deletion of a short segment of amino acids (Craik *et al*, 1983)

This finding led to the proposal of a sliding-junction model the variation in length of surface loops being caused by sliding of splice junctions resulting in an enhanced variability of gene products and thereby speeding up evolution Following this model it would mean that the amino acids 63 and 64 of the αA_2 chain are located at the protein surface and hence that the insertion occurs in a surface region It is, however not clear whether the insertion sequence can be seen as merely an extension of a surface loop It might also be considered as a means to vary the gene products, whereby the evolutionary rate is enlarged The fact that differential splicing has been found in a considerable number of genes (Early *et al* 1980, Young *et al*, 1981, Rosenfeld *et al*, 1983, Schwarzbauer *et al*, 1983, Nabeshima *et al*, 1984, Transy *et al*, 1984) supports the idea that it is a way of trying out a more or less drastic change in a protein

(c) *A reconstruction of the evolution of the αA gene*

The protein sequences of the αA_2 chain that have been shown to be homologous to the heat-shock proteins of *Drosophila* soybean and *Caenorhabditis* are coded for by exon 3 and exon 4 of the αA crystallin gene These homologies provide a strong indication that duplication of an ancestral gene gave rise to both the heat-shock genes and the α crystallins In contrast, the first two exons of the αA crystallin gene show no homology to the small heat-shock proteins Perhaps the first exon was combined with exon 3 and exon 4 after the duplication Therefore, the αA polypeptide can be considered to obey the hypothesis of Gilbert (1978) that new proteins might result by combining structural and functional domains encoded by exons On the other hand, if the first exon was already present in the most primitive ancestral gene, it must have been lost in the ancestor of the heat-shock genes before the duplications that gave rise to the small heat-shock protein family Because αA_2 and αB_2 crystallin are highly homologous also in their first 60 residues, the observed duplication within the first exon of the αA gene must have taken place before the duplication that resulted in separate αA_2 and αB_2 genes In the evolution of the αA crystallin gene the occurrence of the insertion in *Cricetidae* and *Muridae* can be considered as the most recent major evolutionary event A tentative model for the evolution of the αA crystallin genes that summarizes the major evolutionary events is depicted in Figure 8 From this model it can be seen that drastic changes at the nucleotide level play an important role in the evolution of the αA gene

We thank Mr Leon van den Broek for his work on constructing the genomic library, Dr Yvonne Quax-Jeuken for making the αA_2 cDNA clone available before publication, and Dr Wilfried de Jong for helpful discussions and critical reading of the manuscript The present investigations have been carried out in part under the auspices of the Netherlands Foundation for Chemical Research (SON) and with financial aid of the Netherlands Organization for the Advancement of Pure Research (ZWO)

References

Barker W C Ketcham L K & Dayhoff M O (1978) *J Mol Evol* **10** 265 271
Berk A J & Sharp P A (1977) *Cell* **12** 721 732
Blake C C F (1978) *Nature (London)* **273** 267
Bloemendal H (1977) *Science* **197** 127 138
Bloemendal H (1981) Editor of *Molecular and Cellular Biology of the Eye Lens* pp 5 1 5 12 Wiley New York
Bloemendal H Cohen L H & de Jong W W (1978) *Interdiscipl Topics Geront* **12** 261–270
Blundell T Lindley P Miller L Moss D Slingsby C Tickle T Tuenell B & Wistow W (1981) *Nature (London)* **289** 771 777
Breathnach R & Chambon P (1981) *Annu Rev Biochem* **50** 349–383
Cohen L H Westerhuis L W de Jong W W & Bloemendal H (1978) *Eur J Biochem* **89** 259 266
Craik C S Rutter W J & Fletterick R (1983) *Science* **220** 1125 1129
de Jong W W Cohen L H Leunissen J A M & Zweers A (1980) *Biochem Biophys Res Commun* **96** 648–655
de Jong W W Zweers A Versteeg A & Nuy-Terwindt E C (1984) *Eur J Biochem* **141** 131 140
Dodemont H J (1984) Ph D thesis Catholic University Nijmegen
Dodemont H J Andreoli P M Moormann R J M Ramaekers F C S Schoenmakers J G G & Bloemendal H (1981) *Proc Nat Acad Sci USA* **78** 5320 5324
Dodgson J B & Engel J D (1983) *J Biol Chem* **258** 4623–4629
Doolittle R F (1981) *Science* **214** 149 159
Driessen H P C Herbrink P Bloemendal H & de Jong W W (1981) *Eur J Biochem* **121** 83 91
Early P Rogers J Davis M Calame K Bond M Wall R & Hood L (1980) *Cell* **20** 313 319
Gil, A & Proudfoot N J (1984) *Nature (London)* **312** 473 474
Gilbert W (1978) *Nature (London)* **271** 501
Goldberg M (1979) Ph D thesis Stanford University Palo Alto, Calif
Harding J J & Dilley K J (1976) *Expt Eye Res* **22** 1 74
Hartenberger J L (1985) In *Multidisciplinary Analysis of Evolutionary Relationships among Rodents* (Luckett W P & Hartenberger J L eds) Plenum New York in the press
Ingolia T D & Craig E (1982) *Proc Nat Acad Sci USA* **79** 2360 2364
Jeffreys A J (1982) In *Genome Evolution* (Dover G A & Flavell R B eds) pp 157–176 Academic Press London
Keim P Heinrikson R L & Fitch W M (1981) *J Mol Biol* **151** 179 197
King C R & Piatigorsky J (1983) *Cell* **32** 707 712
King C R & Piatigorsky J (1984) *J Biol Chem* **259** 1822–1826
King C R Shinohara T & Piatigorsky J (1982) *Science*, **215** 985 987
Lerner M Boyle J Mount S Wolin S & Steitz J A (1980) *Nature (London)* **283** 220 224

McDevitt M A Imperiale M J Ali H & Nevins J R (1984) *Cell* **37** 993 999
McLachlan A D (1971) *J Mol Biol* **61** 409 424
Medford R M Nguyen H T Destree A T Summers E & Nadal Ginard B (1984) *Cell* **38** 409 421
Messing J (1983) *Methods Enzymol* **101** 20 78
Messing J Crea R & Seeburg P (1981) *Nucl Acids Res* **9** 309 321
Miyata T & Hayashida H (1981) *Proc Nat Acad Sci U S A* **78** 5739 5743
Miyata T Hayashida H Kikuno R Hasegawa M Kobayashi M & Koike K (1982) *J Mol Evol* **19** 28 35
Moormann R J M den Dunnen J T Mulleners L Andreoli P M Bloemendal H & Schoenmakers J G G (1983) *J Mol Biol* **171** 353–368
Mount S (1982) *Nucl Acids Res* **10** 459–472
Nabeshima R M Nguyen H T Destree A T Summers E & Nadal Ginard B (1984) *Nature (London)* **308** 333 338
Padgett R A Konarska H M Grabawski P J Hardy S E & Sharp P A (1984) *Science* **225** 898 903
Perler F Efstratiadis A Lomedico P Gilbert W Kolodner R & Dodgson J (1980) *Cell* **20** 555–566
Quax Jeuken Y Janssen C Quax W van den Heuvel R & Bloemendal H (1984) *J Mol Biol* **180** 457 472
Quax Jeuken Y Quax W van Rens G Meera Khan P & Bloemendal H (1985) *Proc Nat Acad Sci U S A* In the press
Rimm D L Horness D Kucera J & Blattner F R (1980) *Gene* **12** 301 309
Rogers J & Wall R (1980) *Proc Nat Acad Sci U S A* **77** 1877 1879
Rosenfeld M G Mermod J J Amara S G Swanson L W Sawchenki P E Rivier J Vale W W & Evans R M (1983) *Nature (London)* **304** 129–133
Ruskin B Krainer A R Maniatis T & Green M R (1984) *Cell* **38** 317–331
Russnak R H Jones D & Peter E (1983) *Nucl Acids Res* **11** 3187 3205
Sanger F Coulson A R Barrel B G Smith A J H & Roe B A (1980) *J Mol Biol* **143** 161–178
Schoffl F Raschke E & Nagao R T (1984) *EMBO J* **1** 2491 2497
Schwarzbauer J E Tamkun J W Lemischka I R & Hynes R O (1983) *Cell* **35** 421 431
Schwartz R M & Dayhoff M D (1978) In *Atlas of Protein Sequence and Structure* (Dayhoff M D ed) vol 5 suppl 3 pp 353 358
Siezen R J & Argos P (1983) *Biochim Biophys Acta* **748** 56–67
Slightom J L Blechl A E & Smithies O (1980) *Cell* **21** 627 638
Southgate R Ayme A & Voellmy R (1983) *J Mol Biol* **165** 35–57
Staden R (1977) *Nucl Acids Res* **4** 4037 4051
Staden R (1982) *Nucl Acids Res* **10** 2951 2961
Tatei K Takemura K Mayeda A Fujiwara Y Tanaka H Ishihama A & Oshima Y (1984) *Proc Nat Acad Sci U S A* **81** 6281–6285
Tautz D & Renz M (1981) *J Mol Biol* **172** 229–235
Tomarev S I Zinovieva R D Dolgilevich S M Krayev A S Skryabin K G & Gause G G (1983) *FEBS Letters* **162** 47 51
Transy C Lalanne J & Kourilsky P (1984) *EMBO J* **3** 2383–2386
van der Ouderaa F J de Jong W W Hilderink A & Bloemendal H (1974) *Eur J Biochem* **49** 157–168
Yasuda K Kondoh H Okazaki K Shimura Y & Okada T S (1982) *Fifth International Congress of Eye Research* Abstract Book pp 38
Young R A Hagenbuchle O & Schibler U (1981) *Cell* **23** 451–458

Edited by S Brenner

CHAPTER 3

THE LENS PROTEIN αA-CRYSTALLIN OF THE BLIND MOLE RAT, SPALAX EHRENBERGI: EVOLUTIONARY CHANGE AND FUNCTIONAL CONSTRAINTS

Wiljan Hendriks, Jack Leunissen, Eviatar Nevo,

Hans Bloemendal & Wilfried W. de Jong

Proc. Natl. Acad. Sci. USA 84 (1987) 5320-5324.

Proc Natl Acad Sci USA
Vol 84, pp 5320–5324, August 1987
Evolution

The lens protein αA-crystallin of the blind mole rat, *Spalax ehrenbergi*: Evolutionary change and functional constraints

(protein evolution/rodent phylogeny/substitution rate/retina/photoperiodicity)

WILJAN HENDRIKS*, JACK LEUNISSEN*, EVIATAR NEVO†, HANS BLOEMENDAL*, AND WILFRIED W. DE JONG*‡

*Department of Biochemistry, University of Nijmegen, P O Box 9101, 6500 HB Nijmegen, The Netherlands, and †Institute of Evolution, University of Haifa Mount Carmel, Haifa 31999 Israel

Communicated by Francisco J Ayala, April 13, 1987 (received for review January 5, 1987)

ABSTRACT **The complete structure of the single-copy αA-crystallin gene of the blind mole rat (*Spalax ehrenbergi*) has been determined in order to elucidate the evolutionary effects of the loss of vision on a lens-specific protein and its gene. The αA-crystallin gene appears to have all the necessary transcriptional and translational signal sequences to be expressed in the rudimentary lens of the mole rat and gives rise to probably two protein products by means of alternative splicing, as in rodents with normal vision. Comparisons of the blind mole rat αA-crystallin sequence with αA sequences from other rodents reveal a considerable acceleration of the substitution rate at nonsynonymous positions in the mole rat lineage, which reflects a relaxation of selective constraints, but the acceleration is not to the extent that might be expected if the gene were now without any function. The remaining evolutionary constraints still imposed upon the mole rat αA-crystallin gene may possibly reflect the need for α-crystallin expression as an indispensable component in the developmental program of the atrophied eye.**

Functional constraints working at the protein level are recognized as a major determinant of the rate of molecular evolution (1–3). Although several well-known examples clearly seem to illustrate this principle, such as the fast-evolving fibrinopeptides versus the conserved histones (4), the structure–function relationships of most proteins are, in fact, not understood in sufficient detail to reliably correlate functional constraint with rate of evolution This difficulty may also hamper the clear distinction of other types of constraints, such as amino acid composition (5), transcriptional and translational requirements (6), and developmental programs (7, 8). The existence and importance of these additional selective forces might be revealed by studying genes whose protein products have lost their normal principal function in the course of evolution. The substitution rate of such a gene would be expected to increase and might eventually become as fast as in pseudogenes unless other constraints interfere. A unique opportunity for such a study is offered by the rudimentary eyes of blind vertebrate species The eye-specific genes of these animals are no longer subject to the selective constraints associated with the maintenance of vision

The eye lens protein α-crystallin is very suitable for this purpose because much comparative and evolutionary data are available (9–11) α-Crystallin is a slowly evolving protein (10), and, thus, relaxation of constraints should be readily detectable. An important advantage, too, is the fact that the αA and αB subunits, of which α-crystallin is composed, are both encoded by single-copy genes located on different chromosomes (12) The αA and αB chains have ≈57% sequence homology, due to an ancient gene duplication (10), and conspicuous sequence similarity with the small heat shock proteins indicates that the ancestral α-crystallin gene originated from this protein family (13, 14) αA-Crystallin DNA sequences are mainly available from rodents (14–17), which makes the αA gene of the completely blind mole rat, *Spalax ehrenbergi* (18), an excellent target for a study of evolutionary constraints The eyes of this rodent are highly degenerated (19), as an adaptation to a subterranean way of life probably more than 25 million years (myr) ago (20)

Previous studies have shown that crystallin cDNA clones hybridize with genomic DNA of the mole rat (21) The hybridization signals were comparable to those in rodent species with normal vision In addition, immunofluorescence reactions on sections of the mole rat eye revealed that crystallins are still expressed in the atrophied lens cells In the present study, we have determined the nucleotide sequence of the mole rat αA-crystallin gene and have reconstructed the evolutionary changes in this gene and its product in the mole rat lineage

MATERIALS AND METHODS

Protein Sequence Analysis. Beaver lenses (*Castor canadensis*) were supplied by J E Storer and D Baron (Regina, Canada) and were sent to Nijmegen at ambient temperature in a saturated solution of guanidine hydrochloride, from which the crystallins were reconstituted by controlled dialysis Gundi lenses (*Ctenodactylus gundi*) were collected in Tunisia by J J Beintema and were stored frozen until used. Frozen squirrel lenses (*Sciurus vulgaris*) were provided by J. Bouten (Venlo, The Netherlands) All subsequent procedures have been described in detail elsewhere (refs 9 and 11 and references therein)

Computer Programs. DNA sequence gel readings were recorded, compared, edited, and assembled by the programs of Staden (22) Evolutionary trees were constructed using the programs FITCH, DNAPARS, and DNACOMP, as supplied in the phylogeny inference package PHYLIP, distributed by J Felsenstein. FITCH uses distance matrices to construct trees, without allowing negative branch lengths, DNAPARS infers a phylogeny by minimizing the number of substitutions needed to convert one sequence into another, and DNACOMP constructs trees by maximizing the number of compatible sites in each tree topology To calculate the number of synonymous and nonsynonymous substitutions per site from homologous sequences (23), corrected for multiple events (24), a program was written in C and run on a VAX 11/780 computer

The publication costs of this article were defrayed in part by page charge payment This article must therefore be hereby marked "*advertisement*" in accordance with 18 U S C §1734 solely to indicate this fact

Abbreviations Myr, million years, kb, kilobase(s).

‡To whom correspondence should be addressed

RESULTS

The Structure of the Mole Rat αA-Crystallin Gene. A mole rat genomic library (25) was screened using a rat αA-crystallin cDNA clone (pRLαA-1, ref 15) Four cosmid clones bearing sequences homologous to αA-crystallin were isolated Restriction enzyme digestion followed by gel electrophoresis, blotting, and screening with the rat αA cDNA probe demonstrated that all hybridizing bands in the mole rat genomic DNA that were observed in earlier studies (21) were also present in the cloned fragments The above results confirm that all genomic αA-crystallin sequences were obtained Restriction enzyme mapping revealed that the inserts represent a region of ≈55 kilobases (kb), containing a 7 3-kb *Eco*RI–*Hin*dIII fragment responsible for the hybridization signal with the rat cDNA probe This *Eco*RI–*Hin*dIII fragment was subcloned, clone pSCr-1, which was used for sequence determination, was generated

By using the shotgun strategy of Deininger (26) and the dideoxy sequencing method (27), the nucleotide sequence of the major part of pSCr-1 was established (Figs 1 and 2) This clone contains the complete single-copy αA-crystallin gene, flanked by 2 3 kb of upstream and 0 8 kb of downstream sequences As in mouse (17) and hamster (14), the mole rat αA-crystallin gene consists of four exons and three introns and appears to have all the necessary transcriptional and translational signal sequences In mouse and hamster, the single-copy αA-crystallin gene encodes two primary gene products, αA- and αA^{Ins}-crystallin (14, 17) The minor αA^{Ins} chain has hitherto only been found in muroid rodents It contains an insertion of 23 amino acid residues as compared with normal αA This insertion peptide is encoded by the optional exon II, which is flanked at its 3′ end by the dinucleotide GC instead of the consensus splice donor sequence GT This deviation has been proposed to explain the alternative splicing of the αA transcript (17) The fact that in the mole rat αA gene the 3′ splice junction sequence of exon II is also GC (Fig 2) strongly suggests that the transcripts of this gene also use alternative splicing pathways, which leads to the synthesis of both αA and αA^{Ins} chains

The mole rat insert peptide has three or four amino acid replacements as compared with the insertion sequence of hamster, rat, and mouse αA^{Ins}, whereas the latter three differ from each other at only a single position (Fig 2) This reflects a considerable change in this part of the mole rat protein Because of the limited number of known αA^{Ins} sequences, we further focused our attention on the αA-crystallin sequence as deduced from exons I, III, and IV

Amino Acid Replacements in Mole Rat αA-Crystallin. In order to evaluate the amino acid replacements in the mole rat αA chain in the context of evolutionary change of αA-crystallin in rodents with normal vision, we increased the number of six known rodent αA sequences (9, 10, 14–17) by analyzing the primary structure of αA-crystallin for three additional species squirrel, beaver, and gundi (see the legend to Fig 3) Together these nine αA sequences represent the major groups among the ≈32 rodent families (28) From Fig 3, it is immediately obvious that mole rat αA deviates from the general pattern of αA evolution in rodents In *Spalax* αA, there are no less than nine amino acid replacements, which is in sharp contrast with the few replacements in other rodent αA chains Not even a single replacement has occurred in the investigated muroid rodents mouse, rat, gerbil, and hamster (to which superfamily the mole rat also belongs) or in squirrel

As compared with the average rate of change for αA-crystallin in vertebrates (three amino acid replacements per 100 residues in 100 Myr, ref 10), it can be calculated from Fig 3 that in the mole rat lineage this rate increases to 13% per 100 Myr Moreover, four of the inferred replacements in the mole rat lineage (at positions 12, 29, 60, and 163) occur among those 80 residues that are unchanged in all other 72 known αA sequences, ranging from dogfish to man (refs 10, 11, and 32 and this paper) Three replacements occur at variable positions (51, 172, and 173) in the αA sequence, but they involve residues that have not been observed in other species Only two of the replacements have been found earlier phenylalanine-53→leucine in *Rana temporaria* (32) and asparagine-123→serine in hedgehog and alligator (10, 11) Several of the replacements involve radically different residues These findings all clearly indicate an increased tolerance for change in the primary structure of αA in this blind animal

Increased Rate of Change at Nonsynonymous Sites. The availability of αA nucleotide sequences for rat (15), mouse (16, 17), hamster (14), chicken (M Thompson, J Hawkins, and J Piatigorsky, personal communication), and frog (32) made it possible to study in more detail the possible effects of the loss of vision on the evolution of the αA-crystallin gene in the mole rat An analysis of the codons in the mole rat sequence as compared with those in other αA-crystallin genes did not reveal any change in the biased codon usage (data not shown) Using the coding sequences, the corrected number of nonsynonymous substitutions per nonsynonymous site (K_A^c) was calculated for each pairwise comparison of rodent and chicken sequences (Table 1) From this table it

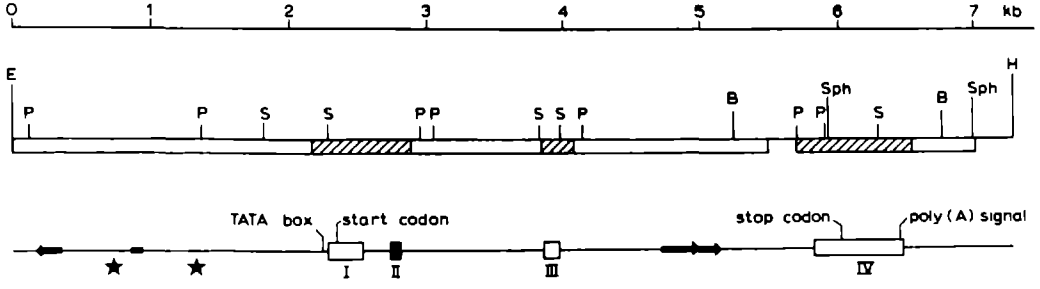


FIG 1 Physical map of the mole rat αA-crystallin gene The second line gives the restriction enzyme cleavage sites that were predicted from the nucleotide sequence and were checked by restriction enzyme analysis E, *Eco*RI, H, *Hin*dIII, S *Sac* I, P, *Pst* I, B, *Bam*HI, Sph, *Sph* I The boxes in the second line represent the two regions for which the nucleotide sequence was established and the hatched parts indicate the sequence regions that are shown in Fig 2 On the third line the lengths and positions of the αA exons are represented by boxes (the optional exon is in black), and the positions of transcription and translation signals are given The asterisks indicate two stretches of alternating purine and pyrimidine bases, $(GT)_{23}$ and $(GT)_{21}$, respectively Three repeated sequences (black arrows) are present that are homologous to the rodent *Alu* type 2 sequences

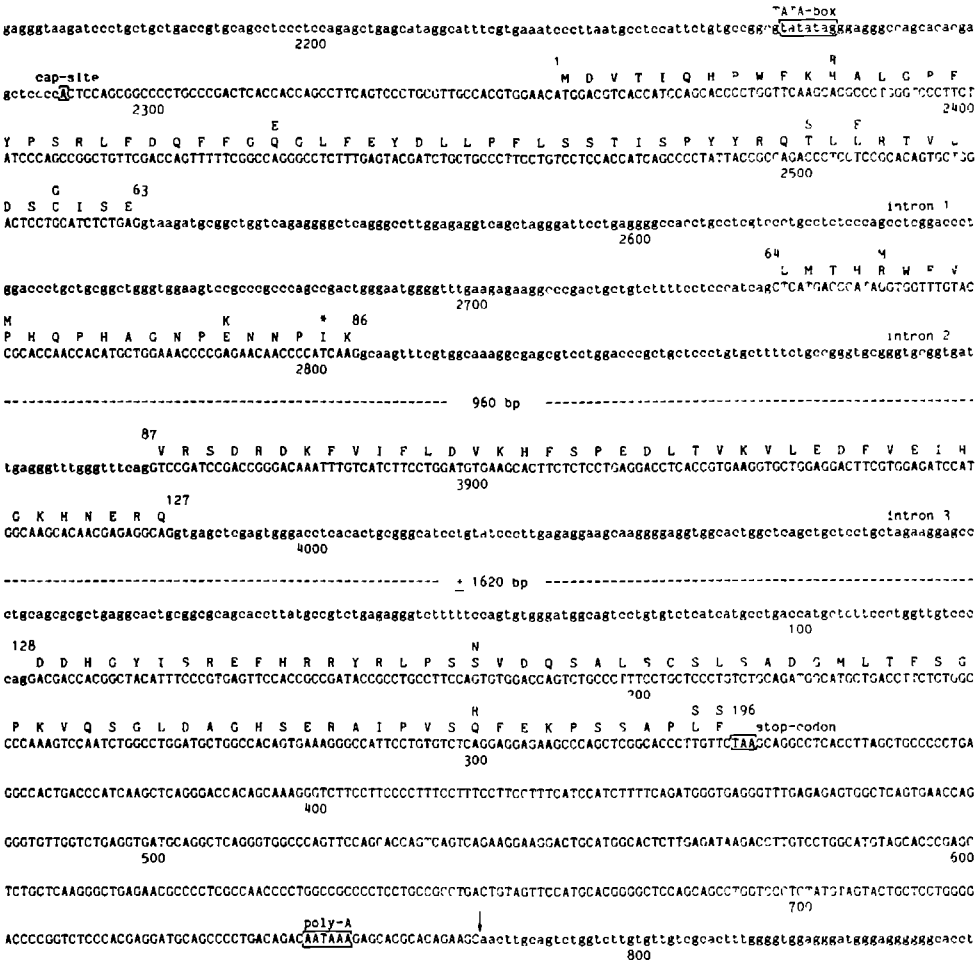


FIG. 2. Sequence of the mole rat αA-crystallin gene. Exon sequences are depicted in capital letters The deduced amino acid sequence is given in single letter notation above the corresponding triplets. Above the mole rat amino acid sequence are the residues that are different in the hamster sequence, which is identical to that of mouse and rat, except for the presence of valine in mouse and glycine in rat (14) at position 85 in the insert peptide (*). The numbering of nucleotides is according to their position in the two sequenced regions of pSCr-1 (Fig 1) The predicted cap site and predicted poly(A) addition site (↓) are indicated The numbering of amino acid residues in this figure includes the optional insert of 23 amino acids (exon II). This gives a discrepancy of 23 in the numbering beyond residue 63 as compared with that in the rest of the paper, where the numbering according to the positions in the normal 173-residue αA chain is adhered to.

appears that in the mole rat lineage 0.023 substitutions per nonsynonymous site have occurred in the 40 Myr since the divergence from the other rodents; this gives an average rate of 0.58×10^{-9} substitution per site per year. This is almost four times the rate in the branches to chicken and the other rodents: 0.15×10^{-9} substitution per site per year (taking the rodent divergence times as given in Fig. 3 and 300 Myr for the divergence of birds and mammals; ref. 33).

The increased rate in the mole rat lineage is a conservative estimate because it is the average over the last 40 Myr of evolution of αA. Although it obviously is uncertain when the effective loss of vision occurred, it seems reasonable to assume that the accelerated evolution of αA only began after the mole rats had fully adapted to a burrowing, subterranean way of life, probably 25 Myr ago according to fossil evidence (20). In this case, the inferred average rate would increase to approximately 0.9×10^{-9} substitution per site per year. This still is only 20% of the rate of change observed in pseudogenes (34), which is assumed to be the neutrality standard for inferring the stringency of selective constraints in sequence regions. Thus, it appears that the αA-crystallin gene in the mole rat has still been under considerable selective constraint, although again it is uncertain whether this constraint has been uniform or discontinuous during the past 25 Myr.

Evolution at Synonymous Sites and in Noncoding Regions. The corrected number of synonymous substitutions per synonymous site (K_S^c) obtained from the comparison of mole rat and chicken αA coding sequences does not significantly differ from the mean value (1.298 with a standard deviation of 0.126) obtained for the other rodent sequences when compared with chicken (Table 1). However, a possible change in the synonymous substitution rate in the mole rat

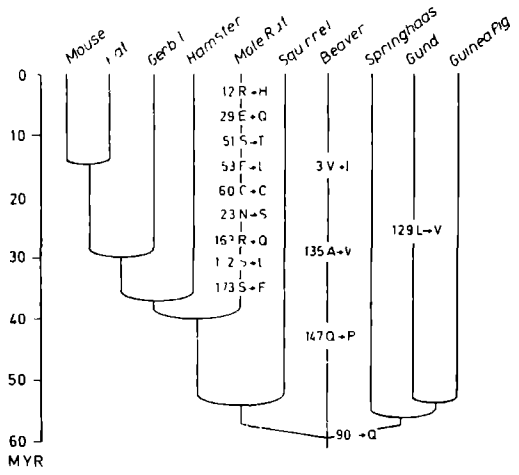


FIG 3 Evolutionary tree showing the inferred distribution and types of replacements in the evolution of αA crystallin in rodents The αA sequences of mole rat squirrel beaver and gundi were added to the already available rodent αA sequences (9 10 14–17) The squirrel beaver and gundi sequences were analyzed by comparing the amino acid compositions of tryptic peptides All peptides were found to have the same compositions as those from rat αA crystallin (9) except for gundi T11 and T17a and also beaver T1 T17b and T18 Subsequent thermolytic digestion of these peptides revealed that the differences were due to the replacements leucine 90→*glutamine and leucine* 129→valine in gundi as well as valine 3→isoleucine alanine 135→valine and glutamine 147→proline in beaver αA crystallin (data not shown) The branching pattern of this tree is based on a current consensus about rodent relationships as inferred from the most recent multidisciplinary evidence (28) apart from the joining of springhaas to the gundi/guinea pig branch which is solely based on the shared replacement leucine 90→glutamine This is however a weak indicator for common ancestry because this replacement and vice versa occur repeatedly in several mammalian orders (10) The divergence time of rat and mouse is averaged from fossil (8–14 Myr ref 29) immunological (22–24 Myr ref 30) and DNA–DNA hybridization data (7–11 Myr F M Catzeflis E N J Ahlquist and G C Sibley unpublished data 17–25 Myr ref 31) The separation time of Gerbillidae from other rodents is according to fossil immunological and hybridization evidence estimated at 19 Myr (29) 35–40 Myr (30) and 36–55 Myr (31) respectively For the divergence time of hamster from the other Muroidea these estimates are 35 Myr (28) 35–40 Myr (30) and 38–58 Myr (31) respectively Because the Spalacidae most probably are the sister group of all other Muroidea (28) their branch should be older than that of the Cricetidae and is placed at 40 Myr in agreement with immunological data (43 Myr ref 30) but earlier than the fossil (25 Myr refs 20 and 28) and hybridization (18–21 Myr F M Catzeflis E N J Ahlquist and G C Sibley *unpublished data*) evidence indicate Immunological data place the origin of squirrels at 55 Myr (30) and the earliest fossil evidence for gundis dates back to 55 Myr (28)

lineage would be hard to detect in this test because the use of the distantly related chicken as an outgroup necessarily results in high and thus less reliable K_S^c values (34) Attempts to deduce such information from comparisons of K_S^c among the rodent αA coding sequences themselves (Table 1) also failed because of the gross uncertainties about rodent divergence times (see legend to Fig 3) Comparisons of the various noncoding regions of the αA crystallin genes of the different species in a search for possible evolutionary effects of the change in selective constraints in the mole rat lineage, were not successful for the very same reasons In fact considerable variation is observed in substitution rates of untranslated regions of mammalian genes, whereas lack of

Table 1 Number of substitutions per site in αA crystallin coding regions

	Mouse	Rat	Hamster	Mole rat	Chicken
Mouse		0 000	0 000	0 023	0 084
Rat	0 227		0 000	0 022	0 095
Hamster	0 293	0 344		0 023	0 085
Mole rat	0 534	0 403	0 410		0 112
Chicken	1 407	1 327	1 160	0 983	

K_A^c (upper right part) and K_S^c (lower left part) values are the numbers of nonsynonymous and synonymous substitutions per site respectively in the αA coding regions corrected for multiple events (23 24)

data have not yet allowed reliable estimates for flanking and intronic regions (34)

Elements that are involved in the expression of the αA crystallin gene have been identified in the 5 flanking sequences of the murine and chicken αA gene (35) Alignment of the first 320 base pairs upstream of the cap site of the mole rat αA gene with the other 5 flanking sequences available (mouse hamster and chicken) revealed that these parts are at least 55% homologous and that the promoter elements are situated in the areas with the best overall homology

Confirmation of Rodent Relationships No mole rat or hamster DNA sequences have yet been used in studies of rodent relationships We therefore used the K_S^c values of Table 1 in a distance matrix method (FITCH) for phylogenetic tree construction and the results showed the mouse and rat αA sequences to be most related followed by hamster and then mole rat The same branching order was consistently obtained when the different αA coding sequences were used in other tree construction methods (DNAPARS and DNACOMP) The phylogenetic information contained in the αA crystallin DNA sequences thus agrees with earlier parsimony analyses including *Spalax* myoglobin and hemoglobin sequences (36) and further strengthens the current opinion about the relationships of Muridae Cricetidae and Spalacidae, as depicted in Fig 3

DISCUSSION

The study described in this paper deals with the fate of the single copy lens-specific αA crystallin gene after loss of its most apparent function—the production of functionally and structurally normal αA crystallin in the transparent lens fiber cells Previous immunofluorescence studies had already indicated that α crystallin is still present in the rudimentary lens of the mole rat (21) The present results show that the mole rat αA gene indeed has all prerequisites for normal expression and gives rise to probably two primary gene products, αA and $αA^{Ins}$ by alternative splicing like in other rodents (14 17) The nonsynonymous substitution rate in the αA crystallin gene has considerably increased in the mole rat lineage but has not reached the value observed in pseudogenes The presence of unexpected more subtle and complex constraints has apparently maintained the gene and its product in the mole rat lineage and kept its average rate of change at a level comparable with for instance that of the globin genes (34)

The fact that the αA gene is still expressed indicates that it is not the gene *per se* that bestows some selective advantage *onto the organism but rather its product One could imagine* that αA crystallin fulfills some as yet unknown role as a minor but essential protein in cells outside the lens Crystallins and their mRNAs have indeed been detected in trace amounts in nonlens tissues (37) In this connection the relation of α crystallin with the small heat shock proteins may also be recalled, although there is no indication that α-

crystallin itself behaves like a heat shock protein or that its gene has a heat shock-inducible promoter.

More likely, perhaps, is the possibility that the expression of the αA gene is an indispensable link in the developmental program of the atrophied eye of the mole rat There is good evidence that this rudimentary eye, though not able to detect light anymore (38), is still of vital importance for photoperiod perception (39), which is required for the physiological adaptations of the animal to seasonal changes It is becoming increasingly clear that in mammals not only the pineal gland is responsible for the induction of such physiological responses (through the synthesis of melatonin directed by photic information received from the retina) but also the retina itself is capable of melatonin synthesis (40) A retinal layer, in which melatonin is synthesized (41), can still clearly be recognized in the mole rat eye (19) If, indeed, this retina is functionally important, it can only be formed through the normal embryological stages of eye development, which necessarily involves the induction of a lens vesicle Crystallin synthesis normally starts after this induction but before the appearance of specific morphological structures (42) It is conceivable that the expression of α-crystallin is an essential step at this stage of eye development It might even be that α-crystallin is more directly involved in the development or function of the retina, considering the finding that Muller glia cells in the neural retina are immunostained by an antiserum to lens antigens enriched for α-crystallin (43)

Our present observations on the evolution of αA-crystallin in the blind mole rat thus lend support to the growing awareness that the concept of "constraints," determining the rate and mode of evolution of a gene, should not be confined to the most obvious functional properties of a gene and its products but should be considered in the much wider context of the complex molecular and cellular interrelationships imposed upon the organism by developmental programs (e g., refs 7 and 8)

We thank M Globulic, Dr F Figueroa, and Prof J Klein for the mole rat genomic library, Dr J E Storer, Dr D Baron, Prof J J Beintema, S Chouref, Ministry of Agriculture (Tunisia), and Jac Bouten BV for rodent lenses, the Israel Discount Bank Chair of Evolutionary Biology and the Ancell-Teicher Research Foundation for financial support, M Versteeg for amino acid analyses Dr T C W Luyben for statistical advice and Dr R van den Heuvel for never-ending inspiration The present investigation was carried out under the auspices of The Netherlands Foundation for Chemical Research and with financial aid from The Netherlands Organization for the Advancement of Pure Research

1 Wilson A C , Carlsson, S S & White, T J (1977) *Annu Rev Biochem* **46,** 573–639
2 Kimura M (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ Press, Cambridge, England)
3 Nei, M & Koehn R K (1983) *Evolution of Genes and Proteins* (Sinauer, Sunderland, MA)
4 Dickerson, R E (1971) *J Mol Evol* **1,** 26–45
5 Graur, D (1985) *J Mol Evol* **22,** 53–62
6 Breathnach, R & Chambon, P (1981) *Annu Rev Biochem* **50,** 349–383
7 Raff, R A & Kaufman, T C (1983) *Embryos, Genes and Evolution* (Macmillan, New York)
8 Maynard Smith, J , Burian, R , Kauffman, S , Alberch, P , Campbell, J , Goodwin, B , Lande, R , Raup, D & Wolpert, L (1985) *Quart Rev Biol* **60,** 265–287
9 de Jong, W W , van der Ouderaa, F J , Versteeg, M , Groenewoud, G , van Amelsvoort, J M & Bloemendal, H (1975) *Eur J Biochem* **53,** 237–242
10 de Jong, W W (1982) in *Macromolecular Sequences in Systematic and Evolutionary Biology*, ed Goodman, M (Plenum, New York), pp 75–114
11 Stapel, S O , Leunissen, J A M , Versteeg, M , Wattel, J & de Jong, W W (1984) *Nature (London)* **311,** 257–259
12 Quax-Jeuken, Y , Quax, W , van Rens, G , Khan, P M & Bloemendal, H (1985) *Proc Natl Acad Sci USA* **82,** 5819–5823
13 Ingolia, T D & Craig, E (1982) *Proc Natl Acad Sci USA* **79,** 2360–2364
14 van den Heuvel, R Hendriks W Quax, W & Bloemendal, H (1985) *J Mol Biol* **185,** 273–284
15 Moormann, R J M , van der Velden, H M W Dodemont, H J , Andreoli, P M , Bloemendal, H & Schoenmakers, J G G (1981) *Nucleic Acids Res* **9,** 4813–4822
16 King, C R , Shinohara, T & Piatigorsky J (1982) *Science* **215,** 985–987
17 King, C R & Piatigorsky J (1983) *Cell* **32,** 707–712
18 Nevo, E (1986) in *Evolutionary Processes and Theory*, eds Karlin, S & Nevo, E (Academic New York), pp 439–474
19 Cei, G (1946) *Monit Zool Ital* **55,** 69–88
20 Klein Hofmeijer G & de Bruijn, H (1985) *Proc K Ned Akad Wet Ser B* **88,** 185–198
21 Quax Jeuken, Y Bruisten, S , Nevo, E , Bloemendal, H & de Jong W W (1985) *Mol Biol Evol* **2,** 279–288
22 Staden, R (1977) *Nucleic Acids Res* **4,** 4037–4051
23 Miyata, T & Yasunaga T (1980) *J Mol Evol* **16,** 23–36
24 Jukes, T H & Cantor C R (1969) in *Mammalian Protein Metabolism*, ed Munro, H N (Academic, New York), pp 21–123
25 Nizetic D , Figueroa, F , Dembic, Z Nevo, E & Klein, J (1987) *Proc Natl Acad Sci USA* **84,** in press
26 Deininger, P L (1983) *Anal Biochem* **129,** 216–223
27 Sanger, F , Coulson A R Barrel, B G , Smith A J & Roe, B A (1980) *J Mol Biol* **143,** 161–178
28 Luckett, W P & Hartenberger, J L , eds (1985) *Evolutionary Relationships Among Rodents A Multidisciplinary Analysis* (Plenum, New York)
29 Jaeger, J J , Tong, H & Denys, C (1986) *C R Acad Sci Ser 2* **302,** 917–921
30 Sarich, V M (1985) in *Evolutionary Relationships Among Rodents A Multidisciplinary Analysis*, eds Luckett W P & Hartenberger J L (Plenum New York) pp 423–452
31 Brownell, E (1983) *Evolution* **37,** 1034–1051
32 Tomarev S I , Zinovieva R D Dolgilevich, S M Krayev, A S , Skryabin K G & Gausse, G G (1983) *FEBS Lett* **162,** 47–51
33 Romer A S (1966) *Vertebrate Paleontology* (University of Chicago Press, Chicago) 3rd Ed
34 Li W -H , Luo, C -C & Wu, C I (1985) in *Molecular Evolutionary Genetics*, ed MacIntyre, R J (Plenum New York) pp 1–94
35 Piatigorsky J (1987) *Invest Ophthalmol Vis Sci* **28,** 9–28
36 Miyamoto M M & Goodman, M (1986) *Syst Zool* **35,** 230–240
37 Clayton, R M Jeanny, J C , Bower, D J & Errington, L H (1986) in *Current Topics in Developmental Biology* eds Okada, T S & Kondoh, H (Academic, Orlando FL) Vol 20, pp 137–152
38 Haim, A , Heth, G , Pratt, H & Nevo, E (1983) *J Exp Biol* **107,** 59–64
39 Pévet, P Heth, G Haim A & Nevo, E (1984) *J Exp Zool* **232,** 41–50
40 Wiechmann, A F (1986) *Exp Eye Res* **42,** 507–527
41 Balemans, M G M Pévet P Legerstee, W C & Nevo, E (1980) *J Neural Transm* **49,** 247–255
42 van Doorenmaalen W J (1981) in *Molecular and Cellular Biology of the Eye Lens*, ed Bloemendal, H (Wiley, New York), pp 415–435
43 Moscona, A A , Fox L , Smith, J & Degenstein, L (1985) *Proc Natl Acad Sci USA* **82,** 5570–5573

CHAPTER 4

MONOCLONAL ANTIBODIES REVEAL EVOLUTIONARY CONSERVATION OF ALTERNATIVE SPLICING OF THE αA-CRYSTALLIN PRIMARY TRANSCRIPT

Wiljan Hendriks, Jo Sanders, Loe de Leij,

Frans Ramaekers, Hans Bloemendal & Wilfried W. de Jong

Eur. J. Biochem. 174 (1988) 133-137.

Eur J Biochem 174 133–137 (1988)

Monoclonal antibodies reveal evolutionary conservation of alternative splicing of the αA-crystallin primary transcript

Wiljan HENDRIKS[1] Jo SANDERS[1], Loe de I FIJ[2], Frans RAMAFKERS[3], Hans BLOEMENDAL[1] and Wilfried W de JONG[1]

[1] Laboratorium voor Biochemie, Universiteit van Nijmegen
[2] Laboratorium voor Klinische Immunologie Universiteit van Groningen
[3] Laboratorium voor Pathologie, St Radboud Ziekenhuis, Nijmegen

(Received December 4, 1987/January 29, 1988) – EJB 87 1365

Because of their specificity and sensitivity, monoclonal antibodies are powerful tools in studies of protein structure and function Therefore, we raised monoclonal antibodies against αA-crystallin and identified the antigenic determinant for two of these antibodies Applying limited-digestion methods, we show that the region spanning residues 158–168 of αA-crystallin contains the epitope for the two monoclonal antibodies These monoclonals were then used to study the occurrence in the lenses of different vertebrates of the elongated αA^{ins}-crystallin chain, a product of alternative splicing It appears that the mutational event resulting in the alternative splicing pattern of the αA-crystallin gene took place at least 70 million years ago This alternative splicing phenomenon has been maintained in rodents and some other, unrelated mammals, but disappeared again in most mammalian lineages

Water-soluble proteins that occur in large amounts in vertebrate lenses are called crystallins About their function little more can be said than that their spatial order probably contributes to the transparency of the ocular lens [1] A remarkable diversity in crystallin composition is observed within and between species, and a variety of biosynthetic mechanisms contribute to the heterogeneity of the different crystallin families [2] One of these mechanisms, alternative splicing, is displayed by the αA-crystallin gene [3, 4]

In several rodents a minor α-crystallin chain (designated αA^{ins}) has been observed that is identical to the normal αA chain of 173 residues, as present in most vertebrates, but which contains an additional sequence of 23 amino acid residues inserted between positions 63 and 64 [3 6] In mouse and hamster it has been shown that the single-copy αA-crystallin gene gives rise to both the normal αA chain and the elongated αA^{ins} chain as a result of alternative splicing of the primary transcript [3, 4] The optional exon of the gene, encoding the insertion, is present only in 10–20% of the mature αA-crystallin mRNA and this ratio remains constant during development [7] The low level of the αA^{ins} mRNA may be caused by the fact that the donor splice junction of the insert exon is AGGC, rather than the consensus AGGT splice sequence [3] The optional exon has also been detected in the αA-crystallin gene of another rodent, the mole rat [8] However, in the chicken αA-crystallin gene no sequence corresponding with the insert exon could be detected [9]

It becomes increasingly evident that alternative RNA processing is a common mechanism to generate multiple products from a single gene (for review see [10, 11]) Little is known, however, about the functional significance of this phenomenon, although in some cases evolutionary conservation of differential splicing pathways (e g [12]) suggests a selective advantage for the organism Moreover, the finding that in many cases the differential splicing processes are regulated in a tissue-specific or a developmental manner seems to imply distinct functions for the different protein products encoded by the same gene (e g [13]) In the case of the αA crystallin gene not only the absence of developmental regulation and the presence of the aberrant GC dinucleotide at the donor splice junction of the optional exon are exceptional, but also the difference between the two products resulting from the alternative splicing a 23-residue segment intruding the structure of a highly conserved protein that was shaped during more than 400 million years of evolution It can hardly be imagined that such an evolutionary 'shotgun' experiment resulted in a functionally advantageous product Therefore, it is of interest to examine the origin of this alternative splicing mechanism and, even more so to reveal why it has been maintained As a first attempt to approach this problem we raised and characterised monoclonal antibodies directed against αA-crystallin and searched for αA^{ins} in lens extracts from a variety of vertebrates

MATERIALS AND METHODS

Preparation of monoclonal antibodies

Two different monoclonal antibodies to αA-crystallin were prepared according to standard procedures [14] Monoclonal Cr A-1 was obtained after immunization of Balb/c mice with bovine lens proteins (five injections, on day 1, 14, 17, 20 and 21, ±100 μg sonified protein per injection) Hybridomas were tested for antibody production and specificity by Western blotting and ELISA techniques Cr A-1 was used as hybridoma culture supernatant and diluted 1 500 for immunoblotting studies Monoclonal antibody Cr I-1,

Correspondence to W W de Jong, Department of Biochemistry, University of Nijmegen, P O Box 9101, NL-6500 HB Nijmegen, The Netherlands

Abbreviation CBB, Coomassie brilliant blue

Enzyme Trypsin (EC 3 4 21 4)

obtained by immunizing mice with purified rat αA^{ins}, was used as ascites fluid and diluted 1:50,000 for blotting experiments.

Limited digestion of α-crystallin subunits

Rat αA- and αA^{ins}-crystallin subunits were isolated as described previously [5]. Partial CNBr cleavage was done according to [5] using reaction times of 2 h in the case of αA, and 1 h for αA^{ins}, at 35 °C. Partial tryptic digestion for 10 min at 30 °C was done as described in [15]. For the CBB-stained gel 2 μg αA, 7.5 μg αA partially digested with trypsin, 6 μg αA partially cleaved with CNBr, 7 μg αA^{ins} partially cleaved with CNBr, and 2 μg αA^{ins} were loaded. For the Western blot the samples were 1, 5, 6, 2 and 0.5 μg, respectively.

Lens extracts

Lenses were obtained from the following rodent species: rat (*Rattus norvegicus*; family Muridae), *Acomys cahirinus* (Muridae), *Psammomys obesus* (Gerbillidae), *Zapus hudsonius* (Zapodidae), squirrel (*Sciurus vulgaris*; Sciuridae), beaver (*Castor canadensis*; Castoridae), *Georichus capensis* (Bathyergidae), guinea pig (*Cavia porcellus*; Caviidae), *Chinchilla laniger* (Chinchillidae), gundi (*Ctenodactylus gundi*; Ctenodactylidae), and *Aplodontia rufa* (Aplodontidae). Representatives of other mammalian orders included hedgehog (*Erinaceus europaeus*; order Insectivora), lemur (*Lemur fulvus*; Primates), rhesus monkey (*Macaca mulatta*; Primates), human (*Homo sapiens*; Primates), rabbit (*Oryctolagus cuniculus*; Lagomorpha), pika (*Ochotona princeps*; Lagomorpha), bat (*Artibeus jamaicensis*; Chiroptera), dog (*Canis familiaris*; Carnivora), ox (*Bos taurus*; Artiodactyla), and whale (*Balaenoptera acutorostrata*; Cetacea). As non-mammalian vertebrates chicken (*Gallus gallus*), emu (*Dromaius novaehollandiae*), alligator (*Alligator mississippiensis*), monitor (*Varanus exanthematicus*), turtle (*Chelydra serpentina*), frog (*Rana esculenta*), cod (*Gadus morhua*), dogfish (*Squalus acanthias*), and lamprey (*Petromyzon marinus*) were studied.

Lenses were homogenized in phosphate-buffered saline (10 mM sodium phosphate, pH 7.2 containing 0.9% sodium chloride), and centrifuged for 15 min at 5000 rpm. The supernatant was dialysed, lyophilized and stored at −20 °C.

Immunoblot analysis

The lens proteins were separated on 15% SDS/polyacrylamide gels and transferred electrophoretically from the gel onto nitrocellulose paper [16]. For a detailed description of the immunoblotting see [17]. Peroxidase-conjugated rabbit anti-mouse immunoglobulin (Dakopatts) was used as second antibody. For the CBB-stained patterns 10 μg lens protein was loaded per lane. For the immunoblots different amounts of total lens protein was used: 1 μg for rat, 2 μg for *Acomys*, *Psammomys*, *Zapus*, beaver, hedgehog, rabbit, pika, dog and ox; 5 μg for *Aplodontia*, squirrel, *Georichus*, guinea pig, chinchilla, gundi, lemur, monkey, human and whale; 15 μg for bat; 30 μg for emu; and 100 μg for chicken, alligator, monitor, turtle, frog, cod, dogfish and lamprey lens extracts. In the case of beaver, pika and bat, isolated α-crystallin was used.

RESULTS AND DISCUSSION

Characterization of monoclonal antibodies

Polyvalent antisera against α-crystallin have proved to be limited in their use due to the complex population of antibodies with a range of binding affinities for each antigenic determinant of the protein and, worse, of the protein's contaminants. Therefore, we prepared monoclonal antibodies directed against αA-crystallin according to standard procedures (see [14]). For the study outlined below monoclonal antibodies Cr.I-1 (raised against rat αA^{ins}-crystallin) and Cr.A-1 (raised against bovine lens proteins) were used. Western blots of two-dimensional gels [18] of rat and bovine lens extracts were made in order to investigate the specificity of these monoclonal antibodies. Both antibodies immunoreacted exclusively with the αA- and αA^{ins}-crystallin chains (data not shown).

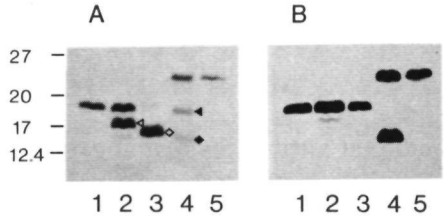


Fig. 1. *Epitope-mapping with monoclonal antibodies against αA-crystallin.* (A) CBB-stained pattern after SDS-PAGE of (1) rat αA, (2) rat αA partially digested with trypsin, (3) rat αA partially cleaved with CNBr, (4) rat αA^{ins} partially cleaved with CNBr, and (5) rat αA^{ins}. (◇) The position of the αA-crystallin fragment containing residues 1 – 138; (◆) the position of the αA^{ins} fragments from residues 66, 69 and 73 to 196; (◀) the αA^{ins} fragment that lacks the C-terminal 35 residues; (◁) the position of the αA fragment of residue 1 – 157. (B) Immunoblotting of the same gel with monoclonal antibody Cr.I-1. Antibody Cr.A-1 gave similar results. Numbers on the left indicate molecular mass in kDa

In order to determine the epitope regions in the αA chains that were recognized by the antibodies, purified rat αA- and αA^{ins}-crystallin were subjected to partial tryptic digestion and to partial CNBr cleavage, followed by electrophoresis and immunoblotting. The results obtained after incubation with the antibodies Cr.I-1 and Cr.A-1 are presented in Fig. 1. Rat αA-crystallin contains methionine residues at positions 1 and 138 [19] (Fig. 2). The long CNBr fragment of αA, missing residues 139 to 173 (Fig. 1 A, lane 3), is clearly not recognized by either monoclonal antibody (Fig. 1 B, lane 3), indicating that the respective epitopes are located in the carboxy-terminal part of the molecule. This is confirmed by the partial CNBr cleavage of rat αA^{ins}. Since the 23-amino-acid insertion peptide contains three additional methionine residues, at positions 65, 68 and 72 in the αA^{ins} chain [5] (and W. H., H. Weetink, H. B. & W. W. de J., unpublished work), fragments from residues 66, 69 or 73 to 196 of the αA^{ins} chain (Fig. 2) are obtained if there is no cleavage after Met-161 (Met-138 in αA-crystallin), and these are immunostained in Fig. 1 B, lane 4. Again, the CNBr fragment missing the C-terminal 35 residues of αA^{ins} is negative. In Fig. 1 A, lane 2, the major fragment generated by trypsin at the primary *in vitro* cleavage site of αA-crystallin (residue 157 [15]), is not recognized by the antibodies (Fig. 1 B, lane 2). Instead, a minor somewhat larger fragment, comprising residues 13 to 173, is immunostained. In addition, on Western blots of human lens extracts that were immunostained using the antibodies Cr.I-1 and Cr.A-1 the predominant *in vivo* degradation product of αA, which misses four or five carboxy-terminal residues [24], is still visible (data not shown). These data map residues 158 to 168 of the αA-crystallin chain as the area containing the epitope for both Cr.I-1 and Cr.A-1.

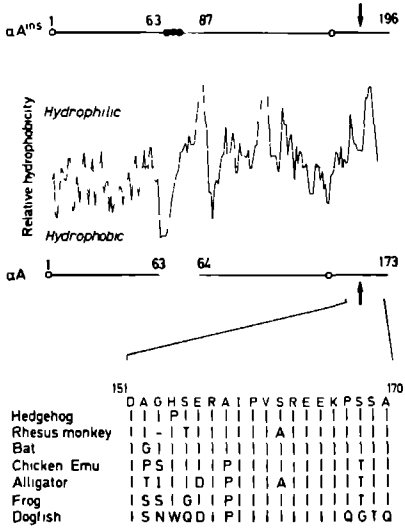


Fig 2 *The epitope region for antibodies Cr I 1 and Cr A-1 in the αA and αA^ins crystallin chain* The upper bar represents the rat αA^ins-crystallin chain The insert peptide is situated between residues 63 and 87 Circles are used to show the positions of methionine residues, the black ones are only present in αA^ins The arrow indicates the primary *in vitro* trypsin cleavage site [15] The lower bar represents the rat αA chain In between, the relative hydrophobicity according to Hopp and Woods [20] for each heptapeptide of rat αA^ins crystallin is plotted against the average sequence position of the heptapeptide In the lower panel the amino acid sequence for residues 151 to 170 of rat αA crystallin is given The corresponding sequences of gerbil squirrel, beaver guinea pig gundi lemur, rabbit pika, dog, ox and whale αA crystallin are identical to that of rat and the replacements and deletion (−) in this region for hedgehog, monkey bat, chicken emu alligator, frog and dogfish αA are shown [8, 19 21−23] No αA-crystallin sequences are available for *Acomys*, *Zapus*, *Georichus* chinchilla, *Aplodontia* monitor, turtle, cod or lamprey

Evolutionary conservation of the epitope region

Previously, on the basis of hydrophilicity profiles, the region containing residues 162−168 has been proposed as one of the most immunogenic epitopes of αA-crystallin [25] Cross-reactivity studies with monoclonal antibodies to bovine α-crystallin were consistent with this prediction [26] Our limited-digestion studies demonstrate that the antigenic determinant for Cr I-1 and Cr A-1 is indeed situated in the carboxy-terminal part of the αA chain In addition, experiments that will be described below show that both monoclonal antibodies react with αA chains from a wide variety of vertebrates, ranging from rat to lamprey (Figs 3−5), demonstrating that this epitope is evolutionary well conserved Comparison of the primary structures of the αA chains used in this study reveals that between positions 156 and 167 only a few replacements occur (Fig 2) Moreover, the differences in affinities of the antibodies observed for αA-crystallin from various sources can be understood from these sequences (Fig 2) For instance, the low affinity of the monoclonal antibodies for non-mammalian αA-crystallin chains may well be ascribed to the replacements 158 Ala→Pro and/or 168 Ser→Thr Neither monoclonal antibody reacted with lens extracts of monitor, cod and dogfish In the case of dogfish this can tentatively be attributed to the 167 Pro→Gln replacement in the predicted determinant Unfortunately, no sequences are available for monitor and cod Taking all these observations together, it seems reasonable to propose that residues 158−168 (Ala-Ile-Pro-Val-Ser-Arg-Glu-Glu-Lys-Pro-Ser) of rat and bovine αA-crystallin contain the antigenic determinant for monoclonal antibodies Cr I-1 and Cr A-1 This is in accord with earlier evidence that residues 146−163 form a surface-exposed segment in the α-crystallin aggregate [15]

Occurrence of αA^ins in vertebrate species

Using both monoclonal antibodies, we surveyed the occurrence of the elongated αA chain, αA^ins, amongst representatives of the approximately 32 rodent families [27] Immunoblotting after SDS-gel electrophoresis of total lens extracts revealed that all investigated rodents display the presence of αA^ins next to the normal αA-crystallin chain, although considerable variation in the ratio of the two chains can be noticed (Fig 3) For a number of species isolated α-crystallin was used, confirming that αA^ins is an integral part of these α-crystallin aggregates

The presence of αA^ins in probably all rodents prompted us to investigate lens extracts from other mammals Results with extracts from lenses of nine different and disparate mammalian species are shown in Fig 4 In hedgehog, bat and pika lenses an elongated αA chain, with the size of αA^ins, is present Again, this chain is an integral subunit of the α-crystallin aggregate of these species as illustrated in Fig 4, lanes 5 and 6 Also human lens extracts both fetal and adult were investigated but no proteins other than the normal αA chain and its primary *in vivo* degradation product could be detected with the two monoclonal antibodies (data not shown)

Changes in the electrophoretic mobilities of αA and of αA^ins can be observed when different species are compared (Figs 3 and 4) It has been shown that differences in SDS gel mobility do not necessarily reflect changes in polypeptide length but may be caused by amino acid replacements that influence the electrophoretic mobility of SDS-protein complexes [28] This also implies that protein sizes as deduced from SDS gel mobility should be interpreted with caution Therefore, it is not surprising to find that the difference in apparent size between αA and αA^ins, as estimated by SDS-gel electrophoresis (4 kDa), is larger than what would be expected from the insertion of a 23-residue peptide

Both monoclonal antibodies were then used to analyse lens extracts of ten sub-mammalian vertebrates for the presence of αA^ins (Fig 5) No cases of elongated αA chains were detected In lens extracts of chicken, emu alligator, turtle, frog and lamprey the normal αA band was still visualized although weaker than in mammals while in monitor, cod and dogfish lens extracts no protein could be immunostained

The use of the appropriate amount of lens extract is of vital importance in these experiments When too much protein is loaded onto the gel non-specific staining occurs, while reduction of the amount of lens extract might result in the immunoreaction being below the limit of detection Although we used a range of sample sizes, we cannot, therefore, be completely sure that the αA^ins chain is altogether absent in species that did not reveal an αA^ins band, especially in the lower vertebrates where a much weaker or even no immunostaining of αA-crystallin occurred

Evolutionary implications

Previously it was believed that αA^ins-crystallin only occurred in the rodent families Muridae and Cricetidae [3−6]

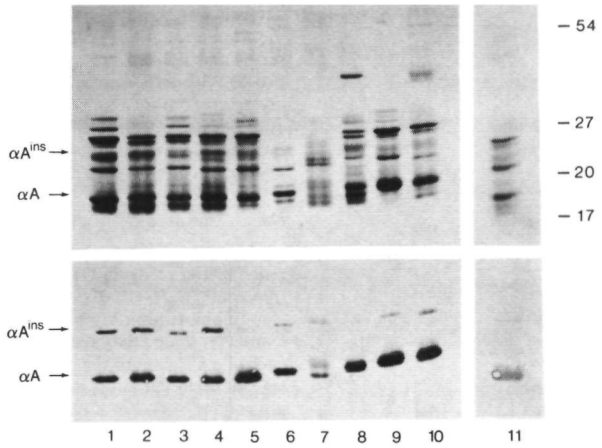


Fig. 3. *The occurrence of αA^ins-crystallin in rodents.* Upper panel: CBB-stained pattern after SDS-PAGE of water-soluble lens proteins of (1) rat, (2) *Acomys*, (3) gerbil, (4) *Zapus*, (5) squirrel, (6) beaver (α-crystallin), (7) *Georichus*, (8) guinea pig, (9) chinchilla, (10) gundi, and (11) *Aplodontia*. Lower panel: immunoblotting of the same gel with the monoclonal antibody Cr.I-1. Antibody Cr.A-1 gave similar results. Numbers on the right indicate molecular mass in kDa

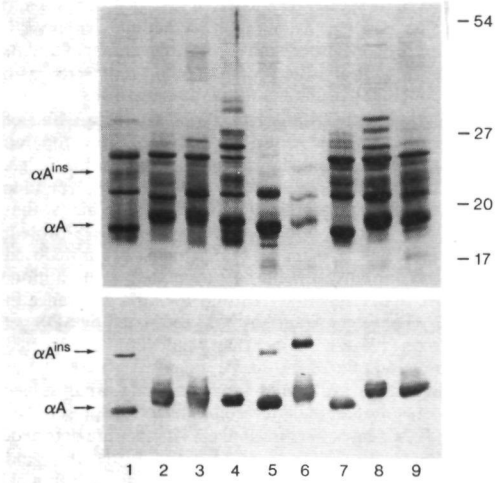


Fig. 4. *The occurrence of αA^ins-crystallin in non-rodent mammals.* Upper panel: CBB-stained pattern after SDS-PAGE of water-soluble lens proteins of (1) hedgehog, (2) lemur, (3) rhesus monkey, (4) rabbit, (5) pika (α-crystallin), (6) bat (α-crystallin), (7) dog, (8) ox, and (9) whale. Lower panel: immunoblotting of the same gel with monoclonal antibody Cr.I-1. Antibody Cr.A-1 gave similar results. Numbers on the right indicate molecular mass in kDa

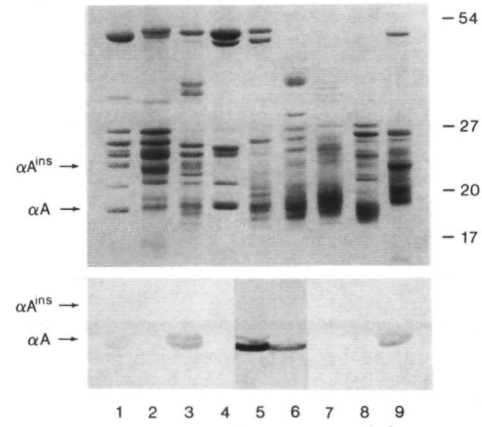


Fig. 5. *The occurrence of αA^ins-crystallin in non-mammalian vertebrates.* Upper panel: CBB-stained pattern after SDS-PAGE of water-soluble lens proteins of (1) emu, (2) chicken, (3) alligator, (4) monitor, (5) turtle, (6) frog, (7) cod, (8) dogfish, and (9) lamprey. Lower panel: immunoblotting of the same gel with monoclonal antibody Cr.I-1. Antibody Cr.A-1 gave similar results. Numbers on the right indicate molecular mass in kDa

Recently it was shown that the αA-crystallin gene of the blind mole rat (family Spalacidae) also contains the 69-bp insert exon [8]. Using immunological techniques we have now traced back further the mutational event that has led to this splicing aberration. Since all investigated rodents, as well as several unrelated mammals like hedgehog, pika and bat, possess the elongated αA^ins chain, the mutational event responsible for the alternative splicing must have occurred before the radiation of the mammalian orders, at least 70 million years ago. This implies, however, that the alternative splicing has independently disappeared again in different lineages, considering the apparent absence of detectable αA^ins in cow, dog, whale, rabbit and primates. In this respect it will be informative to compare the regions corresponding to the optional exon in the αA-crystallin genes of mammals with and without αA^ins.

The question remains, however, why the ability of the αA-crystallin gene to give rise to multiple protein products has been maintained through at least 70 million years of divergent evolution. As mentioned before, the αA^ins subunit is an inte-

gral part of the α-crystallin aggregate Moreover, there are no indications that αA^{ins} is less stable or disturbs the integrity of the rodent α-crystallin aggregate (W H , H Weetink, H B and W W de J , unpublished work) On the other hand, there is no obvious advantage for the occurrence of αA^{ins} next to the normal αA chain If then, at best, the αA^{ins} chain is a selectively neutral character, one wonders why neutral drift has not eliminated this feature in all species or, otherwise, led to 100% splicing of the insert exon in some species For the time being, the reasons for the evolutionary conservation of the alternative splicing of the αA-crystallin gene transcript remain completely obscure

We thank Drs J J Beintema, J Bours, M L Johnson, E Nevo, C G van Zijl de Jong and the Reptile Zoo 'Iguana' for providing eye lenses, John Mulders and Christina Voorter for helpful suggestions and discussions, Gert Schaart and Edwin Schwander for technical assistance and Richard van den Heuvel for never-ending inspiration The present investigation was carried out under the auspices of The Netherlands Foundation for Chemical Research (SON) and with financial aid of The Netherlands Organization for the Advancement of Pure Research (ZWO)

REFERENCES

1 Delaye, M & Tardieu, A (1983) *Nature (Lond) 302*, 415–417
2 de Jong, W W & Hendriks, W (1986) *J Mol Evol 24*, 121–129
3 King, C R & Piatigorsky, J (1983) *Cell 32*, 707–712
4 van den Heuvel, R , Hendriks, W , Quax, W & Bloemendal, H (1985) *J Mol Biol 185*, 273–284
5 Cohen, L H , Westerhuis, L W , de Jong, W W & Bloemendal, H (1978) *Eur J Biochem 89*, 259–266
6 de Jong, W W , Cohen, L H , Leunissen, J A M & Zweers, A (1980) *Biochem Biophys Res Commun 96*, 648–655
7 King, C R & Piatigorsky, J (1984) *J Biol Chem 259*, 1822–1826
8 Hendriks, W , Leunissen, J , Nevo, F , Bloemendal, H & de Jong, W W (1987) *Proc Natl Acad Sci USA 84*, 5320–5324
9 Thompson, M , Hawkins, J & Piatigorsky, J (1987) *Gene 56*, 173–184
10 Leff, S E , Rosenfeld, M G & Evans, R M (1986) *Annu Rev Biochem 55*, 1091–1117
11 Breitbart, R E , Andreadis, A & Nadal-Ginard, B (1987) *Annu Rev Biochem 56*, 467–495
12 Gil, G , Smith, J R , Goldstein, J L & Brown, M S (1987) *Proc Natl Acad Sci USA 84*, 1863–1866
13 Grima, B , Lamouroux, A , Boni, C , Julien, J -F , Javoy-Agid, F & Mallet, J (1987) *Nature (Lond) 326*, 707–711
14 de Leij, L , Poppema, S & The, T H (1983) *J Immunol Methods 62*, 69–72
15 Siezen, R J & Hoenders, H J (1979) *Eur J Biochem 96*, 431–440
16 Towbin, H , Staehelin, T & Gordon, J (1979) *Proc Natl Acad Sci USA 76*, 4350–4354
17 Mulders, J , Hoekman, W , Bloemendal, H & de Jong, W W (1987) *Exp Cell Res 171*, 296–305
18 O'Farrel, P H (1975) *J Biol Chem 250*, 4007–4021
19 de Jong, W W , van der Ouderaa, F J , Versteeg, M , Groenewoud, G , van Amelsvoort, J M & Bloemendal, H (1975) *Eur J Biochem 53*, 237–242
20 Hopp, T D & Woods, K P (1981) *Proc Natl Acad Sci USA 78*, 3824–3828
21 de Jong, W W (1982) in *Macromolecular sequences in systematic and evolutionary biology* (Goodman, M , ed) pp 75–114, Plenum, New York
22 de Jong, W W , Zweers, A , Versteeg, M & Nuy-Terwindt, E C (1984) *Eur J Biochem 141*, 131–140
23 Stapel, S O , Leunissen, J A M , Versteeg, M , Wattel, J & de Jong, W W (1984) *Nature (Lond) 311*, 257–259
24 de Jong, W W , van Kleef, F S M & Bloemendal, H (1974) *Eur J Biochem 48*, 271–276
25 Puri, N , Augusteyn, R C , Owen, E A & Siezen, R J (1983) *Eur J Biochem 134*, 321–326
26 Augusteyn, R C , Seifert, R , Tai, L & Thompson, C H (1985) *Curr Eye Res 4*, 13 -20
27 Luckett, W P & Hartenberger, J L (eds) (1985) *Evolutionary relationships among rodents A multidisciplinary analysis*, Plenum, New York
28 de Jong, W W , Zweers, A & Cohen, L H (1978) *Biochem Biophys Res Commun 82*, 532–539

CHAPTER 5

STRUCTURAL EQUIVALENCE OF SUBUNITS IN THE RAT α-CRYSTALLIN AGGREGATE

Wiljan Hendriks, Hans Weetink, Christina E.M. Voorter,

Jo Sanders, Hans Bloemendal & Wilfried W. de Jong

Submitted for publication.

ABSTRACT

In rodents and some other unrelated mammals, alternative splicing of the αA-crystallin gene transcript results in the synthesis of the elongated αA^{ins}-crystallin chain. This polypeptide is identical to the normal αA-crystallin chain of 173 residues, but contains an additional sequence of 23 amino acid residues inserted between positions 63 and 64. To determine the effects of this insert peptide, the structure of the rat α-crystallin aggregate and its subunits αA-, αA^{ins}-, and αB-crystallin was studied. Fluorescence spectra, partial urea dissociation, and lactoperoxidase-catalyzed iodination of surface residues could not reveal any difference between the αA and αA^{ins} subunits. Moreover, it seems that all α-crystallin subunits occupy equivalent positions in the protein aggregate, and that the insert peptide merely elongates the connecting peptide between the amino- and carboxy-terminal domain of the αA-crystallin subunit.

INTRODUCTION

α-Crystallin, a large multisubunit protein aggregate, is found in high concentrations in the vertebrate ocular lens. Its abundancy is thought to contribute to the transparency and refractive index gradient of the lens, which is essential for the focussing of light onto the retina (1). Two single-copy genes are responsible for the subunits that make up the protein: the αA- and the αB-crystallin genes (2). However, a variety of posttranslational modifications of the two primary gene products make that α-crystallin can be composed of a large number of different polypeptides (3). Although the primary structure of these polypeptides has been studied extensively (e.g. 4) little is known of the spatial conformation of the subunits. α-Crystallin chains mainly contain β-pleated sheets and almost no α-helices (5). Based on several criteria, a two-domain subunit structure has been proposed (6, 7).

How the polypeptides are arranged in the native α-crystallin aggregate is still far from being clear. It has been suggested that the subunits are arranged in three concentric layers, with only A-chains in the core and different proportions of A and B subunits in the other two layers (8). A slightly modified version of this model was proposed recently (9), in which the three-layer model leads to tetrahedral symmetry with 12, 24 and 24 sites in the first, second and third layer, respectively. These two models both implicate that subunits do not occupy equivalent sites. Two other models are

based on the equivalence of subunit positions, and suggest a dodecamer, displaying tetrahedral symmetry (7), and a protein micelle (7, 10), respectively, as the possible quaternary structure for α-crystallin.

In lenses of rodents and some other mammalian species an additional primary gene product next to αA and αB is part of the α-crystallin aggregate (11). It is identical to the normal αA-crystallin chain of 173 residues, as present in most vertebrates, but contains an additional sequence of 23 amino acid residues inserted between positions 63 and 64 (12-16). In mouse and hamster it has been shown that the single-copy αA-crystallin gene gives rise to both the normal αA-crystallin chain and this insert-containing polypeptide, designated αA^{ins}-crystallin, as a result of alternative splicing of the primary transcript (14, 15). The existence and evolutionary maintenance of this aberrant polypeptide raises important questions about the structural characteristics of the α-crystallin subunits. How can a subunit, containing a 23-residue segment intruding the structure of a highly conserved protein, be accommodated as an integral part of the α-crystallin aggregate? To approach this problem we studied tryptophan location, partial dissociation, limited proteolysis, and accessibility of subunits in the rat α-crystallin aggregate. The results imply that αA-, αB- and αA^{ins}-crystallin are structurally alike, and occupy equivalent sites in the rat α-crystallin aggregate.

MATERIALS AND METHODS

Protein isolation. Lenses of two month old Wistar rats were homogenized in 2 volumes of buffer (50 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA, pH 7.6) and centrifugated at 11,000 rpm for 20 min at 5°C. The supernatant was applied onto an Ultrogel AcA 34 column, and peak fractions containing α-crystallin were pooled, dialysed against distilled water, and lyophilized (12). It appeared that lyophilized α-crystallin displayed a considerable spreading in the apparent size for different preparations, as revealed by sedimentation velocity determination. Therefore, when proteins were isolated for conformational studies, this concentration method was not used. Instead, pooled fractions were dialysed against PBS (10 mM sodium phosphate, 0.9% [w/v] NaCl, pH 7.2), concentrated by ultrafiltration through an Amicon UM10 filter, and kept frozen at -20 °C until needed. In this way, rat α-crystallin aggregates with an approximate weight of about 700 kDa were reproducibly obtained. Subunits were isolated by ion-exchange chromatography of rat α-crystallin on a DEAE-cellulose column in the presence of 6 M urea (12).

Sequence determination. Peptides, obtained by tryptic, chymotryptic, or thermolytic digestion, were purified by gel filtration over Sephadex G-50 fine, high-voltage paper electrophoresis, and paper chromatography (12). After hydrolysis in 6 M HCl, amino acid compositions were determined on an LKB Alpha Plus amino acid analyzer. Amino acid sequence determination by automated Edman degradation was performed on an Applied Biosystems gas-phase sequenator.

Fluorescence spectra. Fluorescence spectra were measured on a Perkin Elmer MPF-4 fluorescence spectrophotometer. The excitation monochromator was set to pass 295 nm light with a bandpass of 3 nm. Fluorescence emission was detected at right angles to the excitation beam after passing through the 1.0 cm quartz cell. Samples usually contained 0.5 to 1.0 mg of protein per ml PBS, and were passed through a membrane filter (0.45 µm pores) to reduce scattering.

Urea dissociation. Urea-containing solutions were prepared from an 8 M urea solution that was deionized immediately before use. In addition to various concentrations of urea the buffer (pH 6.3) contained 70 mM sodium phosphate, 80 mM NaCl, 0.04% DTE (w/v). Sephacryl S-200 (Pharmacia) chromatography was performed on a column with dimensions 120 x 1.5 cm at a flow rate of 12 ml/h; 1 ml samples containing 50 mg of protein were applied. Fractions were dialysed against distilled water, lyophilized, and analysed by SDS-gel electrophoresis. Gel scanning was performed on a laser densitometer (LKB 2202 Ultroscan). In the case of iodinated α-crystallin, fractions were monitored for ^{125}I radioactivity in a γ-counter, and the SDS-gel was autoradiographed.

Lactoperoxidase-catalyzed iodination. Iodination was carried out with modifications of the lactoperoxidase method of Thorell & Johansson (17). 125 µg rat α-crystallin and 20 µg lactoperoxidase (Sigma, L-8257) were mixed with 148 µl 0.1 M sodium phosphate, pH 7.5, and 24 µl $Na^{125}I$ (1 mCi, 7 mCi/µg). The reaction was initiated by addition of 1 µl of 500 µg/ml hydrogen peroxide. After 20 min at room temperature sodium bisulphite was added to stop the reaction. To get rid of unincorporated radio-iodine, the iodinated protein was isolated by gel filtration on a 1 x 30 cm Sephadex G-100 column developed with 50 mM sodium phosphate, 0.9% NaCl, pH 7.5. Peak fractions, as monitored with a γ-counter, were combined and kept frozen until needed.

Isolation of iodinated peptides. Iodinated rat α-crystallin was mixed with 200 mg carrier rat α-crystallin, and labeled subunits were isolated as described previously (12). Peak fractions as monitored with a γ-counter were pooled, dialysed, lyophilized, and checked for purity by SDS-gel electrophoresis and subsequent autoradiography. To introduce additional tryptic cleavage sites at cysteine residues, αA and αA^{ins} were S-β-aminoethylated (18). Subunits were digested with trypsin, and subsequent lowering of the pH to 6.5 resulted in the precipitation of some peptides. These core peptides were collected by centrifugation and washing, and were separated by gel filtration over Sephadex G-50 fine (18). In some cases purified core peptides were subsequently digested with thermolysin using established methods (18). Separation of soluble peptides was done with reversed-phase HPLC on a μBondapak C-18 column (19), or micropeptide-mapping at the 1 nmol level (20). Iodinated fractions and spots were monitored with a γ-counter and by autoradiography, respectively. Peptide compositions were determined using an LKB Alpha Plus amino acid analyzer.

RESULTS

The primary structure of the rat α-crystallin subunits. Rat α-crystallin is composed of three types of subunits: αA-, αB- and αA^{ins}-crystallin. The primary structure of the rat αA and αA^{ins} chains have previously been determined (12, 21). Sequence determination of the mouse (14), hamster (15) and mole rat αA genes (16) had already indicated that the published sequence of the insert peptide in rat αA^{ins} (12) probably had to be corrected: two residues should be inverted (positions 77 and 78 in Fig. 1), and a tryptophan should be introduced (position 74 in Fig. 1). Because this tryptophan residue is important with respect to the fluorescence studies described in this paper, we confirmed its presence by determining the sequence of a chymotryptic peptide (position 71 to 75 in Fig. 1) by automated Edman degradation.

To enable the interpretation of conformational studies the amino acid sequence of the rat αB-crystallin chain was determined. Amino acid compositions of tryptic, chymotryptic and thermolytic peptides of rat αB were compared with the hamster αB-crystallin sequence (23). All peptides were found to have the same composition as those predicted from the hamster sequence, except for tryptic fragment T10 (position 117-127; Fig. 1). Subsequent thermolytic digestion of this peptide and comparison with human (24), bovine (22) and dogfish (4) αB sequences revealed that the difference was due

to the replacement 98-Ile → Val in hamster (position 122 in Fig. 1). In Fig. 1 the sequences of the three subunit types of the rat α-crystallin aggregate are shown.

Location of tryptophan residues. Fluorescence spectra of proteins can provide structural information about the relative environment of the tryptophan residues. The fluorescence properties of various tryptophan residues in proteins can be distinguished into three discrete spectral classes (25). The class I tryptophan residues are buried in hydrophobic regions of the protein,

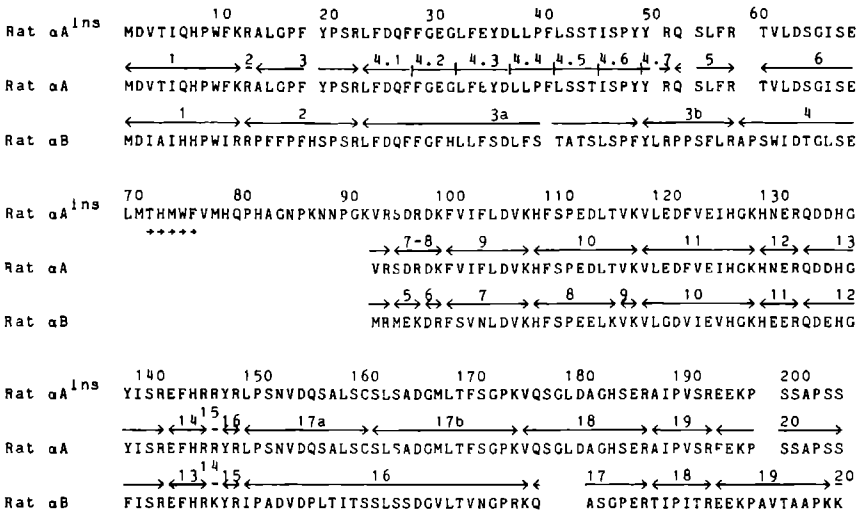


Figure 1. Amino acid sequences for the three different subunits of rat α-crystallin, aligned according to ref. 4. Tryptic peptides are indicated and numbered in accordance with earlier papers (12, 18, 21, 22). Peptides 3a and 3b result from chymotrypsin-like cleavage during the tryptic digestion of αB. An additional tryptic cleavage site at cysteine (position 159), resulting in peptides 17a and 17b, was generated by S-β-aminoethylation of αA and αA^{ins}. Peptides generated by thermolytic cleavage of tryptic peptide 4 of the acidic chains (αA and αA^{ins}) are numbered 4.1, 4.2, etc. The chymotryptic pentapeptide isolated from αA^{ins}-crystallin that was subjected to automated Edman degradation is shown (→).

Table 1. Fluorescence emission maxima (± 1 nm) for rat α-crystallin and its subunit homopolymers. Spectra were determined in two different solvents for at least four independent isolates. The excitation wavelength was set at 295 nm. For comparison the values for L-tryptophan are included.

Compound	λ_m (nm) in PBS	λ_m (nm) in 6 M Urea
Rat α-crystallin	332	352
Rat αA homopolymers	331	351
Rat αA^{ins} homopolymers	332	351
Rat αB homopolymers	334	352
L-Tryptophan	353	354

and have λmax = 330 - 332 nm. Class II residues are in limited contact with water at the protein surface, and have λmax = 340 -342 nm. Class III tryptophans are completely exposed to water, and have λmax = 350 - 353 nm. The relative number of residues in each class varies from one protein to another, and the observed fluorescence spectrum will be a weighted sum of contributions from each class. Because the insert peptide of rat αA^{ins}-crystallin has an additional tryptophan residue next to the one also present in the normal αA-chain (positions 74 and 9, respectively, in Fig.1), the determination of the fluorescence spectra of rat α-crystallin and of the subunit homopolymers could perhaps elucidate the localization of the extra 23 residues in αA^{ins}. In Table 1 the results of these fluorescence measurements are presented. From these data it can be concluded that all rat α-crystallin trypthophan residues are predominantly in hydrophobic environments. Also the additional tryptophan in the insertion peptide is buried in hydrophobic parts of the molecule.

Partial dissociation studies on rat α-crystallin. Dissociation studies on bovine α-crystallin have provided some clues about the architecture of the aggregate (7, 8, 26, 27). Therefore, the dissociation behaviour of rat α-crystallin, containing the elongated αA^{ins} chain, was studied. Rat α-crystal-

lin was applied onto a Sephacryl S-200 column at different urea concentrations, and the results for urea concentrations of 3.0, 4.0, and 5.0 M are shown in Fig. 2. The elution profile at 8.0 M urea revealed, of course, one peak since the aggregate is totally dissociated. Only the first part of this peak contained the αA^{ins} chain, as a consequence of its higher molecular weight. The same pattern was obtained when 6.0 M urea was used. Using no urea does, of course, result in one fast moving protein component: the native aggregate. At a urea concentration of 3.0 M, protein components that move slower through the gel filtration column than the main peak can be observed

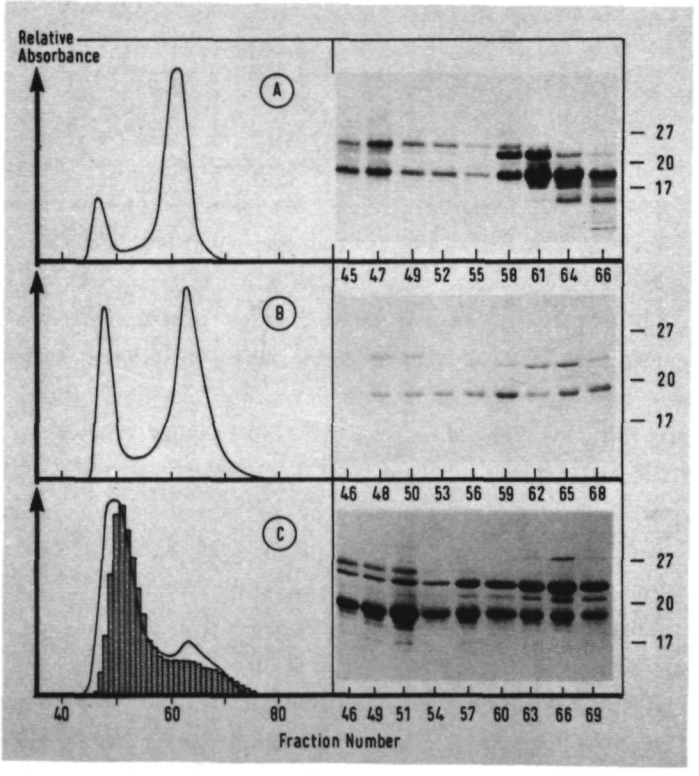


Figure 2. Sephacryl S-200 chromatography of rat α-crystallin at different urea concentrations. Left: elution pattern monitored at 280 nm. Right: CBB-stained pattern after SDS-PAGE of the indicated fractions in the elution pattern. A, 5.0 M urea; B, 4.0 M urea; C, 3.0 M urea. In the case of 3.0 M urea, ^{125}I-labeled α-crystallin was applied, and shaded bars represent the relative radioactivity of the different fractions.

(Fig. 2C). These components, representing subunits that were relatively weakly bound to the aggregate under these conditions, mainly consist of αB chains. Noteworthy is that, although some αA-crystallin chains did, no αA^{ins} chains left the fastest component. At 4.0 M urea, two peaks that are well separated can be observed (Fig. 2B). The second peak, again, contains exclusively αA- and αB-crystallin chains; the first one is comprised of αA and αA^{ins} next to a trace of αB. Moreover, the relative amount of αA^{ins} versus αA has drastically increased in this first peak as compared to the pattern obtained with 3.0 M urea. This finding becomes even more pronounced when 5.0 M urea was used: the first peak was found to contain αA^{ins} and αA only, and in a one to one ratio. Further increase of the urea concentration did not result in a change in this ratio, but led to the disappearance of this first peak. Thus, even under rather denaturing conditions, an interaction between αA and αA^{ins} is present.

Iodination of surface tyrosines by lactoperoxidase. To investigate the relative positions of the three different subunits in the rat α-crystallin aggregate, lactoperoxidase was used to iodinate residues on the surface of the protein. Lactoperoxidase-catalyzed iodination has proven to be a powerful technique for detecting proteins, or parts of proteins, exposed on the surface of supramolecular structures. Since the relatively large size of lactoperoxidase (78 kDa) precludes its penetration into the target structure, no tyrosines in the interior of a protein, sterically hindered, or in an unfavourable micro-environment will be modified (28).

Rat α-crystallin was radio-iodinated by lactoperoxidase. Subsequent gel electrophoresis and autoradiography revealed that the acidic α chains (αA and αA^{ins}) are labeled to more or less the same extent, but a relatively small percentage is incorporated in the αB chains (Table 2). This may be due to the difference in tyrosine content between the acidic and basic chains (Fig. 1). A small amount of iodinated α-crystallin was added to 50 mg unlabeled rat α-crystallin and applied onto the Sephadex S-200 column in the presence of 3 M urea. Fig. 2C shows that, taking into account the relative activity of the three subunit types, the distribution of label throughout the elution profile simply corresponds with the relative prevalence of the subunit types in the different fractions. This demonstrates that the gel filtration experiment does not mimic a "peeling off" of the aggregate, and therefore one cannot argue from the previous experiments that αA^{ins} resides only in the core of the aggregate. In fact, αA^{ins} seems to be just as accessible for iodination

Table 2. Relative iodine content of tryptic peptides, isolated from the different rat α-crystallin subunits. The nomenclature of the peptides is according to the literature, and is indicated in Fig. 1. Residues are numbered in accordance with their appearance in the primary structure, and in brackets their position in the alignment of Fig. 1 is indicated to facilitate intersubunit comparisons.

Subunit	% of total protein	% of iodine content	Fraction	% label of subunit	Possible peptide	Possible residue
αA	65.0	87.8	A1	2.3	16	Tyr-118 (146)
			A2	61.2	13	Tyr-109 (137)
			A3	3.9	13	Tyr-109 (137)
			A4	7.0	3	Tyr- 18 (19)
			A5	1.2	4.7	Tyr- 48 (49)
			A6	17.3	4.6	Tyr- 47 (48)
			A7	1.2	4.6	Tyr- 47 (48)
			A8	5.9	4.3	Tyr- 34 (35)
αA^{ins}	6.3	7.4	I1	4.0	16	Tyr-141 (146)
			I2	52.9	13	Tyr-132 (137)
			I3	5.6	13	Tyr-132 (137)
			I4	12.0	3	Tyr- 18 (19)
			I5	1.6	4.7	Tyr- 48 (49)
			I6	13.4	4.6	Tyr- 47 (48)
			I7	3.5	4.6	Tyr- 47 (48)
			I8	7.0	4.3	Tyr- 34 (35)
αB	28.7	4.8	B0	19.5	3	Tyr- 48 (49)
			B1	4.0	15	Tyr-122 (146)
			B2	6.4	13	His-119 (143)
			B3	12.1	10	His-101 (125)
			B4	58.0	3b	Tyr- 48 (49)

by lactoperoxidase as αA (Table 2). In addition, limited digestion studies on rat α-crystallin showed that αA^{ins} and αA are equally susceptible to trypsin (data not shown).

Isolation of labeled peptides. Although the data in Table 2 show that αA and αA^{ins} are iodinated to more or less the same extent, it might well be that the distribution of label over various residues in both chains is different due to the insertion peptide. Therefore, to localize the iodinated sites in the different α chains, subunits were isolated from labeled rat α-crystallin and digested with trypsin. The resulting peptides were separated by reversed-phase HPLC and micropeptide-mapping, and incorporated radioactiv-

ity was assigned to peptides obtained from the subunits.

For a number of reasons it turned out to be rather complicated to correlate labeled fractions in the chromatographic experiments and labeled spots on the 2D-micropeptide maps with α-crystallin subunit peptides. Due to the incorporation of iodine, the retention times of tyrosine-containing peptides on reversed phase HPLC are increased, and the mobilities on cellulose thin-layer plates are altered as compared to the noniodinated peptide (e.g. 29, 30). Furthermore, depending on the environment, tyrosines can be iodinated once (mono-iodo-tyrosine) or twice (di-iodo-tyrosine) and, although tyrosine residues are a much better substrate for lactoperoxidase, iodination of histidine residues is also possible (28).

Tryptic digestion of S-β-aminoethylated rat αA-crystallin creates over 20 different peptides (Fig. 1). Soluble peptides (all but T4 and T9) were separated on a reversed-phase HPLC system, and by amino acid analysis it was possible to assign absorbance peaks to the different tryptic peptides of αA (Fig. 3A). One major (A2) and three minor radioactive peaks (A1, A3 and A4) could be observed. Tyrosine residues occur in the insoluble peptide T4, and in tryptic peptides T3, T13 and T16. Taking into account that retention times of (di)iodinated peptides are higher than those of the unmodified peptides, it seems reasonable to conclude that peaks A1, A2, and A4 correspond with peptides T16, T13, and T3, respectively. Peak A3 could be explained by microheterogeneity of the major radioactive peak. Microheterogeneity of peptides, resulting in peak multiplicity, is a common phenomenon in chromatography, and can be increased by iodination due to a varied extent of labeling.

To get additional support for these assignments, the individual radioactive peaks were loaded on cellulose thin-layer plates and micropeptide mapping (20) was performed. A tryptic digest of unlabeled rat αA was used as carrier and reference (Fig. 3B). Taking into account that the electrophoretic and chromatographic behaviour of peptides is altered by iodination, one can indeed correlate spot A1 with peptide T16, A2 and A3 with T13, and A4 with T3 (Table 2).

The insoluble core peptides (T4 and T9) that were obtained after tryptic digestion of rat αA, were separated on a Sephadex G-50 column (12). Radioactivity was found only in the fractions containing T4. To be able to correlate the amount of activity to the three different tyrosine residues that are present in T4 (Fig. 1), this peptide was digested with thermolysin and the generated fragments (Fig. 1) were analyzed by reversed-phase HPLC and

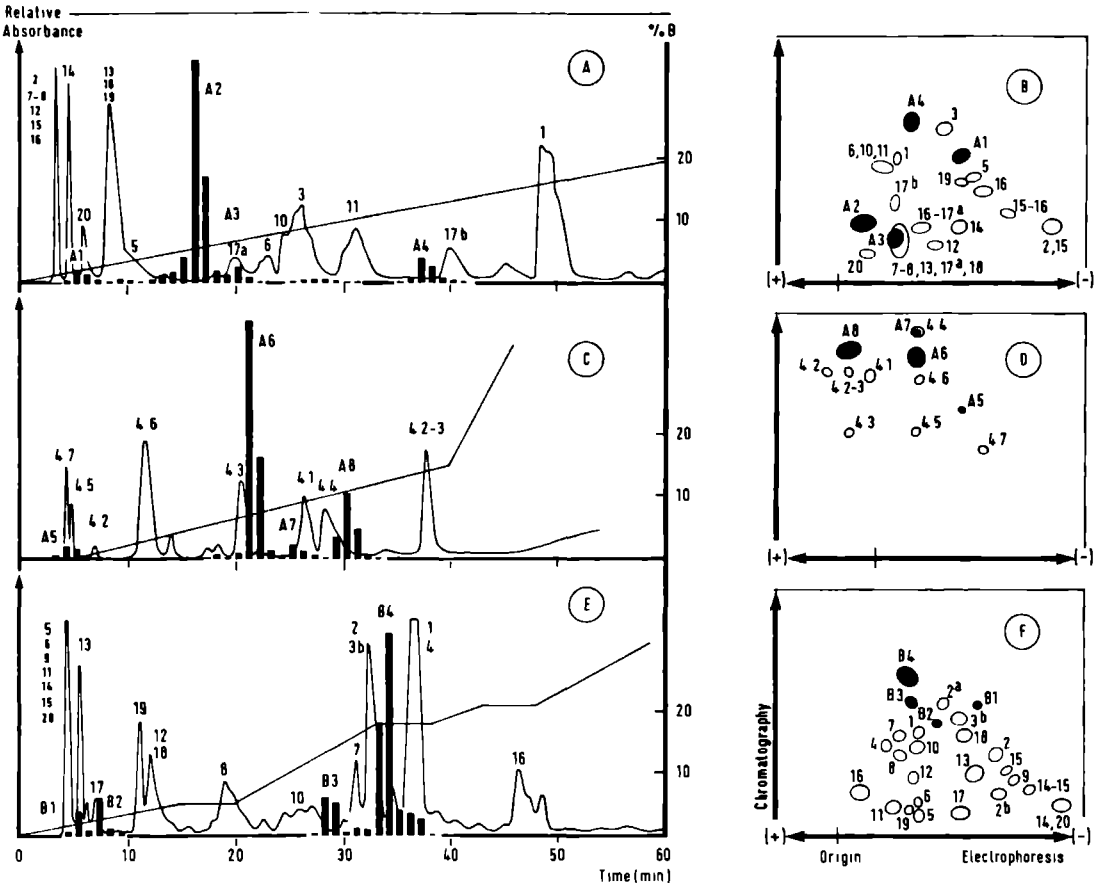


Figure 3. Separation of ^{125}I-containing peptides. Left: elution profiles, monitored at 225 nm, of tryptic peptides from rat αA- (A) and αB-crystallin (E), and of thermolytic peptides from rat αA tryptic fragment 4 (C). The peptides were fractionated by HPLC on a μBondapak C-18 column with a gradient of acetonitrile (0-40%) in 0.1% trifluoroacetic acid, as indicated in the figure (%B). Peptides were identified by amino acid analysis, and are numbered according to their position in the sequence of the protein (cf. Fig. 1). Black bars indicate the relative ^{125}I-content of the corresponding fractions, and radioactive peaks are numbered A1-A8 and B1-B4. Right: the corresponding peptide maps of tryptic peptides from rat αA- (B) and αB-crystallin (F), and of thermolytic peptides from rat αA tryptic fragment 4 (D). Black spots represent the locations of the different radioactive peptides. The correlation between the spots on the map and the radioactive peaks in the elution pattern was established by making peptide maps of individual peaks.

micropeptide mapping (Fig. 3C and 3D). Using arguments parallel to the ones above, the major radioactive peaks, A6 and A8, can be correlated with peptides T4.6 and T4.3, respectively. The minor ones, A5 and A7, are linked with T4.7 and microheterogeneity in T4.6, respectively. Amino acid analysis revealed that a considerable percentage of peptide T4 was lacking its last two residues (Tyr-Arg) due to a chymotrypsin-like cleavage after Tyr-47 (position 48 in Fig.1). Therefore, one has to bear in mind that the amount of label found in peptide T4.7 in fact underestimates the labeling of Tyr-48, and, as a consequence, the label that has been assigned to T16 (also Tyr-Arg) originates in part from T4.7.

The same strategy was used to identify the iodinated residues in the rat αA^{ins}-crystallin chain. αA^{ins} does not contain additional tyrosines as compared to αA, and apart from the replacement of peptide T6 by two other tryptic fragments (see ref. 12) there is no difference in the elution profiles and peptide maps of the unlabeled αA^{ins}-crystallin carrier protein as compared to αA. Interestingly, also no differences were observed between the radioactive peaks in the HPLC elution profiles of αA and αA^{ins}, and identical radioactive spots on the peptide maps were obtained (data not

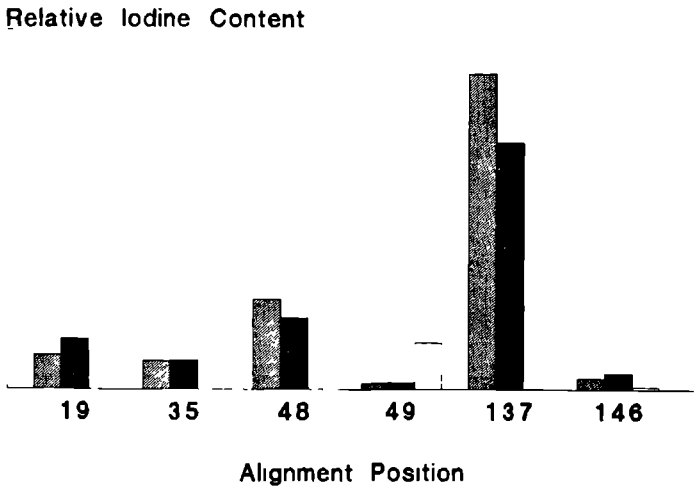


Figure 4. Relative specific activities of the tyrosine residues in αA (hatched bars), αA^{ins} (black bars) and αB (open bars) subunits isolated from the iodinated rat α-crystallin aggregate. The residue numbers refer to the positions in the aligned sequences (Fig. 1).

shown). These results, summarized in Table 2, indicate that the tyrosine residues in αA and αA^{ins} are similarly exposed.

Tryptic digestion of rat αB-crystallin, which contains only two tyrosine residues, results in some 20 peptides (Fig. 1). Insoluble core peptides were again separated over Sephadex G-50, and peptides T3, T3a, and T7 were identified by amino acid analysis. Only T3, containing a single tyrosine residue that corresponds with the fourth tyrosine in the αA chain (position 49 in Fig. 1), appeared to be iodinated. This radioactive peak is referred to as B0 in Table 2. Like for the αA chain, the soluble tryptic peptides of αB were analyzed by reversed-phase HPLC and micropeptide mapping (Figs. 3E and 3F). One major (B4) and three minor (B1, B2, and B3) radioactive peaks were detected. In line with the change in chromatographic behaviour due to iodination, B1 could be assigned to peptide T15 and B4 to T3b. The other two radioactive components are probably peptides containing a iodohistidine residue. For B2 this could be T13, and for B3 perhaps T10.

In Table 2 the quantitative data for the different fractions are given. From this table one can calculate the relative iodine incorporation for each tyrosine in the three different types of subunits by correcting the percentage label of the different tyrosines for the specific activity of the subunit containing these tyrosines. These results are depicted in Fig. 4 and illustrate clearly the similar distribution of label in rat αA- and αA^{ins}-crystallin. It also explaines why αB is labelled so poorly. This subunit does not contain the four tyrosine residues that contribute more than 90% of the incorporated iodine in the αA and αA^{ins} chains.

DISCUSSION

The functional significance and evolutionary persistence of the elongated αA^{ins}-crystallin chain are enigmatic. Knowing that α-crystallin is evolutionarily very conservative, with an extraordinary avoidance of replacements that affect changes in charge (31), it is hard to imagine that the sudden introduction of a 23-residue insert would have no dramatic effects on its structure and function. Nevertheless, the αA^{ins} chain has been part of the α-crystallin aggregate in several mammalian groups, including all rodents, for at least 70 million years (11). Understanding of this insertion phenomenon is hampered by the fact that little is known about the spatial conformation of the α-crystallin subunits, and how they build up the aggregate.

In the present study we have tried to shed some light on the three-dimensional structure of rat α-crystallin and its subunits by comparing data obtained for the three different subunits: αA^{ins}-, αA- and αB-crystallin. The fluorescence emission maxima for rat α-crystallin and the subunit homopolymers demonstrated that all tryptophan residues in α-crystallin are in hydrophobic environments, making it unlikely that the tryptophan-containing insertion peptide in the αA^{ins} subunit is totally on the surface of the molecule. This is supported by tryptic digestion studies showing that first carboxy- and amino-terminal cleavage sites in the αA^{ins} chain are used by trypsin before it cuts at positions in the insert peptide (data not shown). Tryptic digestion of native rat α-crystallin displayed an equal susceptibility for both acidic chains. Again, identification of surface residues by lactoperoxidase catalyzed halogenation again demonstrated equivalence of αA and αA^{ins}. Looking at the intron-exon distribution of the α-crystallin genes (15, 23), and the proposed two-domain structure for α-crystallin subunits (6, 7), it thus seems that the insert peptide is nothing but a connecting peptide between two structural units.

The tyrosine at position 137 is by far the most accessible for iodination (Fig. 4). It is near the position were an intron interrupts the coding sequence of the α-crystallin genes, and is situated very close to the proposed connecting peptide between the two halves of the carboxy-terminal domain (6). From studies on sulfhydryl groups in α-crystallin it is known that in human αA-crystallin Cys-142 (position 170 in Fig. 1) is surface-exposed (32). This is also in agreement with the proposed two-domain structure, because this position would be close to the end of the second structural domain, which then continues into the carboxy-terminal tail. This tail (starting at about position 176 in Fig. 1) indeed contains candidates for the surface region of the subunits like the predominant in vivo and in vitro cleavage sites (33, 34) and the very conserved region spanning an antigenic determinant (11) that is suggested to be important for subunit contacts (33).

From Fig. 4 it is clear that tyrosines in αA and αA^{ins} are equivalent with respect to their solvent accessibility. In addition, positions 49 and 146 are poorly accessible for iodination by lactoperoxidase in both acidic and αB subunits. This, like the finding that any combination of subunits from any species will form viable α-crystallin particles (7), suggests structural equivalence of the different subunits. Of course, this does not mean that the structures are identical. This is illustrated by the partial dissociation experiments. αA^{ins} displays a clear interaction with αA chains, and not with

αB, under denaturing conditions, resulting in a compound containing a one to one ratio of both acidic chains. Probably, the long connecting peptide between the two domains in αA^{ins} interacts with a region that is present in an αA subunit, but not in αB or αA^{ins} itself.

For almost all quaternary structures that contain equivalent subunits, it has been found that such subunits occupy also equivalent positions in the multimeric structure (35). Therefore, it should be no surprise to find that subunits occupy equivalent sites in the α-crystallin aggregate. Indeed, αA and αA^{ins} were found to be equally susceptible to trypsin (data not shown) in the rat α-crystallin aggregate. A comparison between αB and both acidic chains is not possible because of the considerable sequence difference resulting in different cleavage sites. In addition, iodination of surface residues in rat α-crystallin and subsequent partial dissociation and gel filtration (Fig. 2C) suggests that all subunits occupy equivalent positions and/or continuously change places in the aggregate. There is indeed experimental evidence for native α-crystallin molecules to exchange subunits, suggesting a dynamic quaternary structure for the α-crystallin aggregate in which subunit positions are equal (27). Moreover, preliminary studies on iodinated bovine α-crystallin have shown that the phosphorylated and non-phosphorylated chains also occupy equivalent places in the native aggregate (C.V. and W.W. de J., unpublished results).

Our data show that indeed the tertiary structure of the αA^{ins} chain is not seriously affected by the insertion peptide. Moreover, in the rat α-crystallin aggregate the position of this subunit is indistinguishable from that of both other chains. Considering the apparent selectively neutral character of the αA^{ins} subunit, one wonders whether eventually species will be found in which mutations in the αA-crystallin gene have resulted in only one primary gene product for this gene, the αA^{ins}-crystallin chain.

Acknowledgements. We thank Marlies Versteeg for amino acid analyses, Eric Keizer for sedimentation analyses, Martin Salden for cooperation during the iodination experiments, John Mulders and Piet van den Oetelaar for valuable discussions, and Richard van den Heuvel for never-ending inspiration. Amino acid sequences were kindly determined at the SON Gas-Phase Sequenator Facility (State University of Leiden). The present investigation was carried out under the auspices of The Netherlands Foundation for Chemical Research (SON) and with financial aid of the Netherlands Organization for Scientific Research (NWO).

REFERENCES

1 Bloemendal, H. (1981) Molecular and Cellular Biology of the Eye Lens, Wiley, New York.

2 Wistow, G.J. & Piatigorsky, J. (1988) Annu. Rev. Biochem. 57, 479-504.

3 de Jong, W.W., Mulders, J.W.M., Voorter, C.E.M., Berbers, G.A.M., Hoekman, W.A. & Bloemendal, H. (1988) in Advances in Post-translational Modifications of Proteins and Aging (Zappia, V., Galetti, P., Porta, R. & Wold, F., Eds.) Adv. Exp. Med. Biol., Vol. 231, Plenum Press, New York, pp. 95-108.

4 de Jong, W.W., Leunissen, J.A.M., Leenen, P.J.M., Zweers, A. & Versteeg, M. (1988) J. Biol. Chem. 263, 5141-5149.

5 Li, L.K. & Spector, A. (1974) Exp. Eye Res. 19, 49-57.

6 Wistow, G. (1985) FEBS Lett. 181, 1-6.

7 Thomson, J.A. (1985) α-Crystallin: A Model for Multisubunit Protein Assemblies, PhD Thesis, University of Melbourne, Australia.

8 Bindels, J.G., Siezen, R.J. & Hoenders, H.J. (1979) Ophthalmic Res. 11, 441-452.

9 Tardieu, A., Laporte, D., Licinio, P., Krop, B. & Delaye, M. (1986) J. Mol. Biol. 192, 711-724.

10 Augusteyn, R.C. & Koretz, J.F. (1987) FEBS Lett. 222, 1-5.

11 Hendriks, W., Sanders, J., de Leij, L., Ramaekers, F., Bloemendal, H. & de Jong, W.W. (1988) Eur. J. Biochem. 174, 133-137.

12 Cohen, L.H., Westerhuis, L.W., de Jong, W.W. & Bloemendal, H. (1978) Eur. J. Biochem. 89, 259-266.

13 de Jong, W.W., Cohen, L.H., Leunissen, J.A.M. & Zweers, A. (1980) Biochem. Biophys. Res. Commun. 96, 648-655.

14 King, C.R. & Piatigorsky, J. (1983) Cell 32, 707-712.

15 van den Heuvel, R., Hendriks, W., Quax, W. & Bloemendal, H. (1985) J. Mol. Biol. 185, 273-284.

16 Hendriks, W., Leunissen, J., Nevo, E., Bloemendal, H. & de Jong, W.W. (1987) Proc. Natl. Acad. Sci. USA 84, 5320-5324.

17 Thorell, J.I. & Johansson, B.G. (1971) Biochim. Biophys. Acta 251, 363-369.

18 van der Ouderaa, F.J., de Jong, W.W. & Bloemendal, H. (1973) Eur. J. Biochem. 39, 207-222.

19 Voorter, C.E.M., Mulders, J.W.M., Bloemendal, H. & de Jong, W.W. (1986) Eur. J. Biochem. 160, 203-210.

20 Whittaker, R.G. & Moss, B.A. (1981) Anal. Biochem. 110, 56-60.

21 de Jong, W.W., van der Ouderaa, F.J., Versteeg, M., Groenewoud, G., van Amelsvoort, J.M. & Bloemendal, H. (1975) Eur. J. Biochem. 53, 237-242.

22 van der Ouderaa, F.J., de Jong, W.W., Hilderink, A. & Bloemendal, H. (1974) Eur. J. Biochem. 49, 157-168.

23 Quax-Jeuken, Y., Quax, W., van Rens, G., Meera Khan, P. & Bloemendal, H. (1985) Proc. Natl. Acad. Sci. USA 82, 5819-5823.

24 Kramps, J.A., de Man, B.M. & de Jong, W.W. (1977) FEBS Lett. 74, 82-84.

25 Burstein, E.A., Vedenkina, N.S. & Ivkova, M.N. (1973) Photochem. Photobiol. 18, 263-279.

26 Thomson, J.A. & Augusteyn, R.C. (1984) J. Biol. Chem. 259, 4339-4345.

27 van den Oetelaar, P.J.M. (1988) Structural Studies on α-Crystallin and on Racemization of Aspartic Acid Residues, PhD Thesis, University of Nijmegen, The Netherlands.

28 Morrison, M. & Schonbaum, G.R. (1976) Ann. Rev. Biochem. 45, 861-888.

29 Strickland, T.W., Williams, J.F. & Pierce, J.G. (1984) Int. J. Peptide Protein Res. 24, 328-336.

30 Stanton, P.G. & Hearn, M.T.W. (1987) J. Biol. Chem. 262, 1623-1632.

31 de Jong, W.W., Leunissen, J.A.M., Hendriks, W. & Bloemendal, H. (1988) in Molecular Biology of the Eye: Genes, Vision & Ocular Disease (Piatigorsky, J., Zelenka, P. & Shinohara, T., eds.) UCLA Symposia on Molecular and Cellular Biology, new series, Liss, New York, vol. 88, in press.

32 Augusteyn, R.C., Hum, T.P., Putilin, T.P. & Thomson, J.A. (1987) Biochim. Biophys. Acta 915, 132-139.

33 Siezen, R.J. & Hoenders, H.J. (1977) FEBS Lett. 80, 75-80.

34 Siezen, R.J. & Hoenders, H.J. (1979) Eur. J. Biochem. 96, 431-440.

35 Klotz, I.M., Darnall, D.W. & Langerman, N.R. (1975) in The Proteins (Neurath, H., ed.) Vol I, 3rd ed., Acad. Press, New York, pp. 293-411.

CHAPTER 6

DUCK LENS ε-CRYSTALLIN AND LACTATE DEHYDROGENASE B4 ARE IDENTICAL: A SINGLE-COPY GENE PRODUCT WITH TWO DISTINCT FUNCTIONS

Wiljan Hendriks, John W.M. Mulders, Michael A. Bibby,

Christine Slingsby, Hans Bloemendal & Wilfried W. de Jong

Proc. Natl. Acad. Sci. USA 85 (1988) 7114-7118.

Proc Natl Acad Sci USA
Vol 85, pp 7114–7118 October 1988
Biochemistry

Duck lens ε-crystallin and lactate dehydrogenase B_4 are identical: A single-copy gene product with two distinct functions

(glycolytic enzymes/protein evolution/thermostability/three-dimensional molecular models)

WILJAN HENDRIKS*, JOHN W M MULDERS*, MICHAEL A BIBBY†, CHRISTINE SLINGSBY†, HANS BLOEMENDAL*, AND WILFRIED W DE JONG*‡

*Department of Biochemistry University of Nijmegen P O Box 9101 6500 HB Nijmegen The Netherlands and †Department of Crystallography Birkbeck College University of London Malet Street London WC1E 7HX United Kingdom

Communicated by Clement L Markert June 13, 1988

ABSTRACT **To investigate whether or not duck lens ε-crystallin and duck heart lactate dehydrogenase (LDH) B_4 are the product of the same gene, we have isolated and sequenced cDNA clones of duck ε-crystallin. By using these clones we demonstrate that there is a single-copy *Ldh-B* gene in duck and in chicken. In the duck lens this gene is overexpressed, and its product is subject to posttranslational modification. Reconstruction of the evolutionary history of the LDH protein family reveals that the mammalian *Ldh-C* gene most probably originated from an ancestral *Ldh-A* gene and that the amino acid replacement rate in LDH-C is approximately 4 times the rate in LDH-A. Molecular modeling of LDH-B sequences shows that the increased thermostability of the avian tetramer might be explained by mutations that increase the number of ion pairs. Furthermore, the replacement of bulky side chains by glycines on the corners of the duck protein suggests an adaptation to facilitate close packing in the lens.**

The water-soluble lens-specific proteins of vertebrates are called crystallins and can be divided into distinct families (1) In the last few years, considerable knowledge has been gathered about the evolutionary origin of these different types of crystallins α-Crystallin is related to the small heat shock proteins (2–4) and to the major *Schistosoma* egg antigen (5) The β- and γ-crystallins belong to one protein superfamily (6) and have a striking structural similarity with a bacterial spore coat protein (7) Other crystallins have recently been found to be related to enzymes δ-Crystallin, the major protein of the embryonic lens in birds and reptiles, is closely related to argininosuccinate lyase (8, 9) τ-Crystallin, present in lenses of some birds, reptiles, fish, and lampreys, shows strong similarity with human and yeast α-enolase (8), and ρ-crystallin in frog lenses belongs to the same protein superfamily as aldose reductase, aldehyde reductase (10), and prostaglandin F synthase (11)

ε-Crystallin was the first example of the occurrence of an active enzyme as an abundant lens protein (12) It is found in many bird and crocodile lenses at levels of up to 23% of total protein (12, 13) Duck lens ε-crystallin appeared to be almost indistinguishable, on the basis of partial amino acid sequence analysis, from duck heart lactate dehydrogenase (LDH) B_4, and it displayed almost the full activity of this glycolytic enzyme (12) This raised some important molecular evolutionary questions What makes an LDH-B-like protein suitable to function as a structural lens protein? Does this relate to the reported thermal stability of LDH-B in most birds and higher reptiles (14)? Most interestingly, is ε-crystallin the same gene product as authentic heart LDH-B_4, or rather are the two proteins the products of very similar, recently duplicated genes?

MATERIALS AND METHODS

Construction and Screening of the Duck Lens cDNA Library. Total duck (*Anas platyrhynchos*) lens mRNA was isolated by using the lithium/urea procedure (15) Poly(A)-selected mRNA was used to construct an oligo(d T)-primed cDNA library in λgt11 (16) Approximately 250,000 plaques were obtained The library was screened with antiserum against duck ε-crystallin (13) To obtain longer cDNA clones, the library was subsequently rescreened with ^{32}P-labeled 5′ ε-crystallin probes

Sequence Analysis and Comparison. Nucleotide sequences were determined using the DNase shotgun strategy (17) and the dideoxy sequencing method (18) DNA sequence gel readings were recorded, compared, edited, and assembled by the programs of Staden (19) Evolutionary trees were constructed by using the programs FITCH and PROTPARS, as supplied in the phylogeny inference package PHYLIP by J Felsenstein (University of Washington, Seattle)

RESULTS

Primary Structure of Duck Lens ε-Crystallin. Partial amino acid sequence data have previously indicated that duck lens ε-crystallin and duck heart LDH B_4 are very similar, if not identical (12) To determine the complete primary structure of duck lens ε-crystallin, we constructed a λgt11 expression library of duck lens mRNA and searched for ε-crystallin-specific cDNA clones The nucleotide sequences of the cDNA inserts of three positive recombinant clones were determined Since the overlapping sequences of the three cDNA clones are completely identical (Fig 1), we conclude that they represent DNA copies of the same mRNA template This is further corroborated by the fact that all three inserts identified the same transcript of 1 45 kilobases (kb) in duck lens mRNA (Fig 2) The size of this transcript is in excellent agreement with the earlier estimate for the ε-crystallin messenger size (1450 nucleotides, ref 13) The combined sequence,§ as presented in Fig 1, contains 1185 nucleotides and lacks the 5′ noncoding region of the ε-crystallin mRNA

From the coding part of the cDNA, the primary structure of duck lens ε-crystallin was deduced (Fig 1) This sequence starts exactly at the amino-terminal residue, as can be deduced from comparison with known LDH sequences (20, 24–32) All peptides of duck ε-crystallin for which the amino

The publication costs of this article were defrayed in part by page charge payment This article must therefore be hereby marked "*advertisement*" in accordance with 18 U S C §1734 solely to indicate this fact

Abbreviation LDH, lactate dehydrogenase

‡To whom reprint requests should be addressed

§The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur Mol Biol Lab Heidelberg) (accession no J03869)

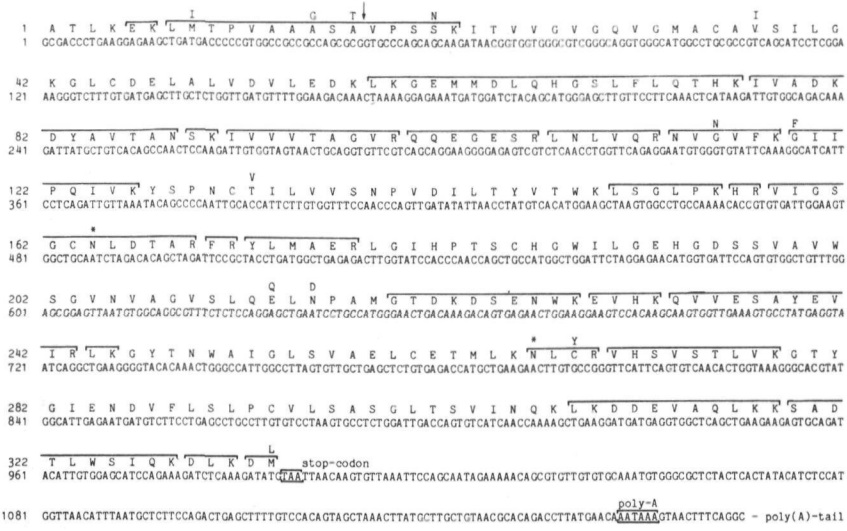


FIG. 1. Combined nucleotide sequences of the cDNA inserts of λDLε-5, λDLε-16, and λDLε-42 and the deduced amino acid sequence of duck ε-crystallin. The λDLε-5 and λDLε-16 inserts span the region from position 592 and 98, respectively, to the poly(A) tail. The λDLε-42 insert contains the first 452 bases. The stop codon and the poly(A)-addition signal are indicated. Above the amino acid sequence are the residues that are different in chicken LDH-B (20). Tryptic peptides whose amino acid compositions were determined in an earlier study are indicated by overbars, and asparagine residues that are probably deamidated in ε-crystallin (12) are marked by asterisks. The numbering of residues is according to ref. 21 in order to facilitate comparison of LDH sequences. As a consequence, residue 16 is followed by residue 18 (arrow), which reflects the presence of an additional residue in the human and porcine sequence.

acid compositions were determined previously (12) could be identified (Fig. 1). Comparative peptide mapping of duck ε-crystallin and duck heart LDH-B_4 had already revealed that the only differences were changes in the charge of two peptides (12). This was suggested to be due to *in vivo* deamidation of asparagine residues 164 and 266 (Fig. 1) of LDH-B in the lens (12). The corresponding codons in the cDNA sequence indeed reveal asparagine residues at these positions in the primary gene product (Fig. 1). This apparently confirms the deamidation of ε-crystallin in the duck lens, but it could also be taken as evidence that we actually have isolated LDH-B instead of ε-crystallin cDNA clones. In the latter case, at least two closely related *Ldh-B*-like genes should be present in the duck genome.

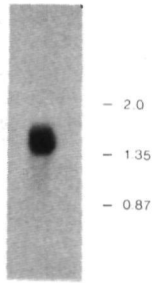


FIG. 2. Size of duck lens ε-crystallin/LDH-B mRNA. Ten micrograms of poly(A)-selected mRNA was glyoxylated and size-fractioned on 1% agarose gels and transferred to nitrocellulose membranes (22). Hybridization was according to Church and Gilbert (23) with the addition of 100 μg of herring sperm single-stranded DNA per ml of hybridization mixture. Washes were done twice at 65°C for 30 min with 0.5 M $NaHPO_4$, pH 7.2/1 mM EDTA/1% $NaDodSO_4$. Sizes (in kb) are indicated at right.

Genomic Organization of ε-Crystallin/LDH-B Sequences. To determine the complexity of genomic ε-crystallin/LDH-B sequences, Southern blots of restriction enzyme digests of duck genomic DNA were hybridized with 5′ and 3′ ε-crystallin probes (Fig. 3). Because in the chicken lens no ε-crystallin is detectable (12, 13), the DNA of this species was used for comparison. From Fig. 3*A* it is obvious that hybridization with the 5′ probe results in only a single band in all chicken digests. In addition, the chicken 1.9-kb *Hin*dIII fragment that is detected with the 5′ probe is absent in the pattern obtained by using the 3′ probe, which in turn hybridizes with fragments of 3.0 and 0.95 kb (Fig. 3*B*). These findings clearly indicate that in the chicken genome there is only one ε-crystallin/LDH-B related sequence: the single-copy functional *Ldh-B* gene.

Southern blots of digested duck DNA, screened with the 5′ probe, reveal one predominant band for each digest (Fig. 3*A*). In the *Eco*RI digest, for example, this is a 4.4-kb fragment. In contrast, the 3′ probe hybridizes with a 19-kb and a 2.0-kb *Eco*RI fragment (Fig. 3*B*). It thus appears that, as in the chicken, there is only one ε-crystallin/LDH-B-related gene present in the duck genome. This implies that duck ε-crystallin and LDH-B_4 are the product of the same single-copy functional *Ldh-B* gene. The observed charge differences between duck heart LDH-B_4 and duck lens LDH-B_4 (ε-crystallin) must therefore indeed be due to posttranslational modifications.

In addition to the predominant fragment in each lane, screening of the digests of duck DNA with the 5′ probe reveals a very faint second band (Fig. 3*A*). These signals are not the result of cross-hybridization with *Ldh-A* or *Ldh-C* gene sequences because in that case there should be similar signals in the chicken digests too. In fact, additional blotting experiments with the λDLε-42 insert (see Fig. 1) as a probe demonstrated that these faint bands represent sequences in the 5′ part of the duck *Ldh-B* gene. This provides evidence

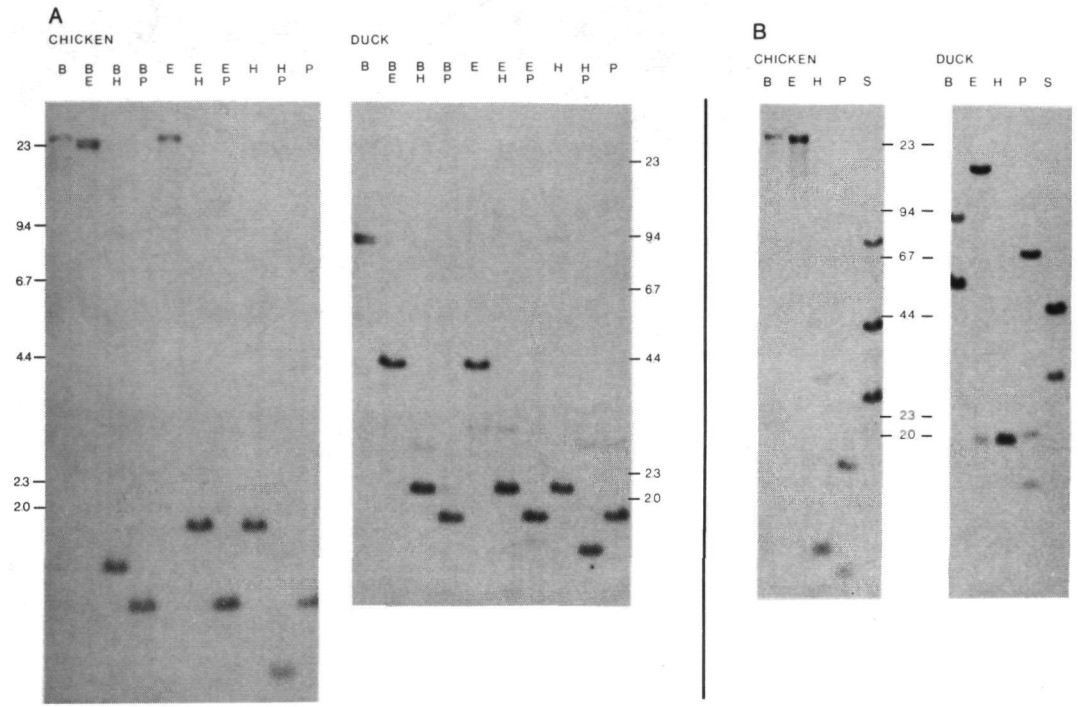


FIG. 3. Detection of ε-crystallin/LDH-B sequences in chicken and duck DNA. Ten micrograms of high molecular weight chromosomal DNA was completely digested with different restriction enzymes, size-fractioned on 0.7% agarose gels, and transferred to nitrocellulose membranes (22). Hybridizations and washes were as described in Fig. 2. (*A*) A 5′ probe (a DNA fragment from nucleotide 98 to the *Sau*3A site at position 187; Fig. 1) was used. (*B*) The λDLε-5 insert (Fig. 1) was used as a probe for 3′ sequences. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst* I; S, *Sac* I. Sizes (in kb) are indicated.

that, as in the *Ldh-A* gene (28, 29), the codons for residues 42 and 43 are separated by an intervening sequence in the *Ldh-B* gene.

Evolution of the *Ldh* Gene Family. Since duck ε-crystallin is in fact duck LDH-B_4, the sequence presented in Fig. 1 can be compared with other known LDH sequences (20, 24–32) to reconstruct the evolutionary history of this protein family. LDH subunits can be divided into three distinct but related classes, encoded by the *Ldh-A*, *Ldh-B*, and *Ldh-C* genes, respectively, which must have originated by duplication from an ancestral gene in the course of evolution (33). There is no evidence that the *Ldh-C* genes of fishes and those of mammals and birds are orthologous; they could be the result of at least two independent duplication events (31, 34). Comparison of the untranslated parts of the different mRNA types reveals that the length of the 3′ noncoding region of the LDH-B mRNAs (some 200 nucleotides; ref. 32) corresponds more with that of mammalian LDH-C mRNA (≈110 bases; refs. 30 and 31) than with that of LDH-A mRNA, which is considerably longer (up to 500 nucleotides; refs. 28 and 29). Whereas the 3′ noncoding region of human and duck LDH-B mRNA displays 56% sequence identity, no significant sequence similarity can be detected between these regions of the three different LDH-type mRNAs. In other gene families too, the length of the 3′ untranslated region has been found to be conserved in orthologous but not in paralogous mRNAs (e.g., ref. 35). This has been taken as evidence for a role of this region in the regulation of expression at the transcriptional or posttranscriptional level.

Three evolutionary scenarios have been proposed for the relationship between the mammalian *Ldh* genes. The mammalian *Ldh-C* gene could have originated from a duplication of an *Ldh-A* gene (31, 36); it might have been derived from a duplicated *Ldh-B* gene (33); or it was already present before the divergence of the *Ldh-A* and *Ldh-B* genes (21). Together with 12 other LDH sequences, the duck LDH-B sequence was used in different phylogenetic reconstruction programs. The result is depicted in Fig. 4 and shows that the testis-specific mammalian *Ldh-C* genes most probably originated by duplication of an ancestral *Ldh-A* gene before birds and mammals diverged. This finding suggests that the same duplication event may also have led to the *Ldh-C* genes in birds, as suggested by Markert *et al.* (33).

Interestingly, the branch lengths in Fig. 4 reveal that since the divergence of rodents and primates the rate of amino acid replacements in LDH-C is about four times the rate in LDH-A. The pronounced acceleration in the rate of change for LDH-C was first suggested by Whitt (36) and was recently supported by immunological and genetic data provided by Millan *et al.* (31). It clearly indicates a change in evolutionary constraints for this type of LDH, in agreement with its testis-specific occurrence.

Structural Changes in Duck LDH-B. LDH forms active tetramers, and the crystallographic structure of pig heart LDH-B_4 has been determined (37). Molecular modeling was used to investigate whether the nine positions where the duck sequence differs from human, pig, and chicken LDH-B might reflect adaptations to the additional function in the lens, which requires close packing ability (38). There are five replacements at subunit contacts (positions 8, 14, 21, 37, and 268), but they would not appear to affect the structure of the LDH-B tetramer. Three changes (residues 115, 119, and 333)

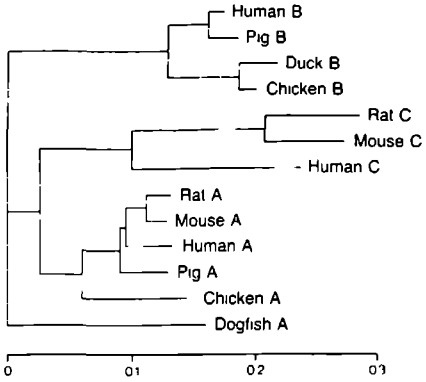


FIG 4 Phylogenetic tree of the LDH protein family This unrooted tree was constructed by using the program FITCH and the scale for branch lengths is in minimum mutational distances per amino acid residue No other topologies could be obtained without allowing negative branch lengths References to the sequences are given in the text The program PROTPARS revealed two equally parsimonious trees both requiring a minimum of 510 amino acid replacements this tree and the one with the chicken LDH A branch and the LDH-C lineage exchanged Maximum parsimony thus, is unable to resolve the trichotomy of the branches to mammalian LDH-C, mammalian LDH A, and chicken LDH A All possible alternative topologies, however, require at least five additional replacements

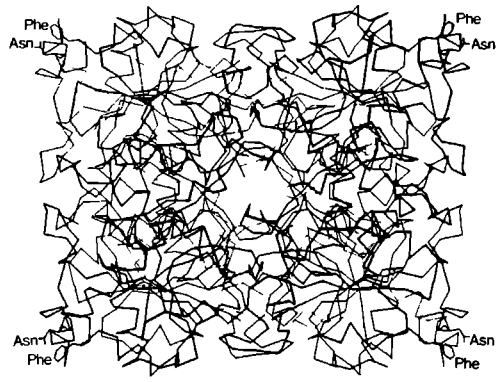


FIG 5 α-Carbon diagram of the three-dimensional model of the pig LDH B_4 tetramer with the side chains of Asn-115 and Phe-119 appended on each subunit In this view, Asn-115 and Phe 119 are seen to occupy prominent surface positions in the LDH-B_4 tetramer, whereas in the duck sequence these have been replaced by glycines The pig LDH B structure was investigated by using the P5LDH coordinates (resolution 2 7 Å) in the Brookhaven Data Bank [Brookhaven National Laboratory (1980) Brookhaven Protein Data Bank (Brookhaven Nat l Lab Upton NY) File P5LDH BRK] and the programs FRODO (40) and MIDAS (I J Tickle, personal communication) on the Evans and Sutherland picture system II Accessibilities of side chains were estimated by running the Lee and Richards program (41) on both the monomers and the dimers subsequently generated This also indicated the side chains that are involved in subunit contacts

occur on the surface of the tetramer The carboxyl-terminal Leu-333 in the pig tetramer is replaced by methionine in duck It has been suggested that the large number of surface methionines found in the oligomeric β crystallin family could act as subunit contact sites, based on the observation that these side chains are frequently involved in crystal lattice interactions in the monomeric γ-crystallins (39) Noteworthy is the radical replacement by glycines of the bulky residues 115 and 119 on the most protruding regions of the tetramer (Fig 5) Finally, change of a crevice position (residue 132) does not appear to perturb the structure None of these changes involve the active site

Replacements common to both chicken and duck LDH-B may indicate structural changes responsible for the increased stability of the avian proteins (14), an advantageous condition for recruiting this enzyme as a lens structural protein Increased hydrogen-bonding patterns (42), increased numbers of ion pairs (43), and decreased sulfur-containing residues (44) have all been implicated in contributing towards elevated thermal stability of proteins A number of replacements in the avian tetramers indeed affect charged residues and thus render the surface more basic and increase ionic interactions Moreover, three surface methionines have been eliminated Methionine is prone to oxidation, which may be a signal for protein turnover (45) Although duck LDH-B subsequently has acquired two additional methionines, one of them (residue 8) is removed from solvent by contributing to a subunit contact, and Met-333 is suggested to be involved in tetramer packing interactions There are two additional cysteines, at positions 45 and 260 in the thermophilic proteins, but these are buried in the core of the protein and are not near other cysteines or aromatic side chains In addition, the duck has gained a cysteine (residue 268) at a subunit interface There are mutations that affect subunit interactions also, but these are more difficult to assess

DISCUSSION

The study described in this paper reveals that in the duck as well as in chicken there is a single-copy *Ldh-B* gene In the duck lens, this gene is overexpressed to the extent that it codes for 10% of the total soluble protein (13), which is far beyond any enzymatic need Mutations in duck LDH-B must thus have been selected not only for maintaining dehydrogenase activity of the enzyme but also to enable its occurrence as an abundant lens protein This novel constraint demands structural stability and the ability for close packing in the lens cell cytoplasm (38) Model building of the duck sequence to porcine LDH-B shows that the most striking changes are those at the extremities of the tetramer, where bulky side chains have been replaced by glycines in the duck structural protein (Fig 5) This would suggest that "rounding off" *the corners of the aggregate may be an adaptation to* facilitate close packing in the lens

The occurrence of LDH B and other enzyme-like proteins (1) in the lens is a surprising example of evolutionary flexibility In the case of LDH-B, mutational events in the common archosaurian ancestry of birds and crocodiles could have introduced, by chance, lens-specific promoter elements in the *Ldh B* gene, or perhaps LDH regulatory genes (34) may have changed the normal expression pattern Changes in isozyme tissue patterns are well-known phenomena in the LDH protein family and can be used successfully in the study of the evolution of taxonomic groups (46) The molecular mechanisms responsible for these species differences in isozyme tissue expressions are usually not known Therefore, an investigation of the molecular basis of the appearance of LDH-B_4 in some bird and crocodile lenses would not only contribute to the understanding of lens-specific gene expression (1) but would also provide insight into the regulation of isozyme expression and gene regulatory evolution in its broadest sense

LDH B is a capricious protein it is a ubiquitous and important 'housekeeping" enzyme, but at the same time it can be absent in humans without apparent ill-effects (47) It reaches levels of 5% of total protein in mouse oocytes (34) or

even 23% in some avian lenses (12) Moreover, beside its glycolytic role, it displays unexpected properties like binding to lipid layers (48), and it perhaps may even function as a centrosomal protein (49)

For the time being, it remains unclear whether the haphazardous occurrence of LDH-B in avian and crocodilian lenses is due to selectively advantageous properties or rather represents a dispensable, neutral character that is tolerated because of the suitable stability and packing properties of the enzyme

We thank Dr Jack Schalken for excellent technical advice and Dr Richard van den Heuvel for never-ending inspiration The help of Jack Leunissen, as well as the use of the services and facilities of the Dutch CAOS/CAMM Center, under Grants SON 11-20-700 and STW-NCH-44 0703, are gratefully acknowledged The present investigation was carried out under the auspices of The Netherlands Foundation for Chemical Research (SON) and with financial aid from The Netherlands Organization for Scientific Research (NWO)

1 Wistow, G J & Piatigorsky, J (1988) *Annu Rev Biochem* **57**, 479–504
2 Ingolia, T D & Craig, E (1982) *Proc Natl Acad Sci USA* **79**, 2360–2364
3 van den Heuvel, R , Hendriks, W , Quax, W & Bloemendal, H (1985) *J Mol Biol* **185**, 273–284
4 de Jong, W W , Leunissen, J A M , Leenen, P J M , Zweers, A & Versteeg, M (1988) *J Biol Chem* **263**, 5141–5149
5 Nene, V , Dunne, D W , Johnson, K S , Taylor, D W & Cordingley, J S (1986) *Mol Biochem Parasitol* **21**, 179–188
6 Driessen H P C , Herbrink, P , Bloemendal, H & de Jong, W W (1981) *Eur J Biochem* **121**, 83–91
7 Wistow, G , Summers, L & Blundell, T (1985) *Nature (London)* **316**, 771–773
8 Wistow, G & Piatigorsky, J (1987) *Science* **236**, 1554–1556
9 Piatigorsky, J , O'Brien, W E , Norman, B L , Kalumuck, K , Wistow, G J , Borras, T , Nickerson, J M & Wawrousek, E F (1988) *Proc Natl Acad Sci USA* **85**, 3479–3483
10 Carper, D , Nishimura, C , Shinohara, T , Dietzchold, B , Wistow, G , Craft, C , Kador, P & Kinoshita, J H (1987) *FEBS Lett* **220**, 209–213
11 Watanabe, K , Fujii, Y , Nakayama, K , Ohkubo, H , Kuramitsu, S , Kagamiyama, H , Nakanishi, S & Hayashi, O (1988) *Proc Natl Acad Sci USA* **85**, 11–15
12 Wistow, G , Mulders, J W M & de Jong, W W (1987) *Nature (London)* **326**, 622–624
13 Stapel, S O , Zweers, A , Dodemont, H J , Kan, J H & de Jong, W W (1985) *Eur J Biochem* **147**, 129–136
14 Wilson, A C , Kaplan, N O , Levine, L , Pesce, A , Reichlin, M & Allison, W S (1964) *Fed Proc Fed Am Soc Exp Biol* **23**, 1258–1266
15 Auffray, C & Rougeon, F (1980) *Eur J Biochem* **107**, 303–314
16 Huynh, T V , Young, R A & Davis, R W (1985) in *DNA Cloning Techniques A Practical Approach*, ed Glover, D (IRL, Oxford), pp 49–78
17 Lin, H -C , Lei, S -P & Wilcox, G (1985) *Anal Biochem* **147**, 114–119
18 Sanger, F , Coulson, A R , Barrel, B G , Smith, A J & Roe, B A (1980) *J Mol Biol* **143**, 161-178
19 Staden, R (1977) *Nucleic Acids Res* **4**, 4037–4051
20 Torff, H J , Becker, D & Schwarzwalder, J (1977) in *Pyridine Nucleotide Dependent Dehydrogenases*, ed Sund, H (De Gruyter, Berlin), pp 31–42
21 Li S S -L , Fitch, W M , Pan, Y -C E & Sharief, F S (1983) *J Biol Chem* **258**, 7029–7032
22 Maniatis, T , Fritsch, E F & Sambrook, J (1982) *Molecular Cloning A Laboratory Manual* (Cold Spring Harbor Lab , Cold Spring Harbor, NY)
23 Church, G M & Gilbert, W (1984) *Proc Natl Acad Sci USA* **81**, 1991–1995
24 Eventoff, W , Rossmann, M G Taylor, S S , Torff, H -J , Meyer, H , Keil, W & Kiltz, H -H (1977) *Proc Natl Acad Sci USA* **74**, 2677–2681
25 Kiltz, H H , Keil, W , Griesbach, M Petry, K & Meyer, H (1977) *Hoppe-Seylers Z Physiol Chem* **358**, 123–127
26 Pan, Y -C E , Sharief, F S , Okabe, M , Huang, S & Li, S S -L (1983) *J Biol Chem* **258**, 7005–7016
27 Matrisian, L M , Rautmann, G , Magun, B E & Breathnach, R (1985) *Nucleic Acids Res* **13**, 711–726
28 Li, S S L , Tiano, H F , Fukasawa, K M , Yagi, K , Shimizu, M , Sharief F S , Nakashima, Y & Pan, Y -C E (1985) *Eur J Biochem* **149**, 215–225
29 Chung F Z , Tsujibo, H , Bhattacharyya, U , Sharief, F S & Li, S S -L (1985) *Biochem J* **231**, 537–541
30 Sakai, I , Sharief, F S & Li, S S -L (1987) *Biochem J* **242**, 619–622
31 Millan, J S , Driscoll, C E , LeVan, K M & Goldberg, E (1987) *Proc Natl Acad Sci USA* **84**, 5311–5315
32 Sakai, I , Sharief, F S , Pan, Y -C E & Li, S S -L (1987) *Biochem J* **248**, 933–936
33 Markert, C L , Shaklee, J B & Whitt, G S (1975) *Science* **189**, 102–114
34 Whitt, G S (1984) *Cell Biochem Function* **2**, 134–139
35 Yaffe, D , Nudel, U , Mayer, Y & Neuman, S (1985) *Nucleic Acids Res* **13**, 3723–3737
36 Whitt G S (1983) in *Isozymes Current Topics in Biological and Medical Research*, eds Rattazzi, M C , Scandalios, J G & Whitt, G S (Liss New York), Vol 10, pp 1–40
37 Grau, U M , Trommer, W E & Rossmann, M G (1981) *J Mol Biol* **151**, 289–307
38 Delaye, M & Tardieu, A (1983) *Nature (London)* **302**, 415–417
39 Slingsby, C , Driessen, H P C , Mahadevan, D , Bax, B & Blundell, T L (1988) *Exp Eye Res* **46**, 375–403
40 Jones, T A (1982) in *Crystallographic Computing*, ed Sayre, D (Oxford Univ Press, Oxford), pp 303–317
41 Lee, B & Richards, F M (1971) *J Mol Biol* **55**, 379–400
42 Alber, T , Dao-pin, S , Wilson, K , Wozniak, J A , Cook, S P & Matthews, B W (1987) *Nature (London)* **330**, 41–46
43 Barlow, D J & Thornton, J M (1983) *J Mol Biol* **168**, 867–885
44 Perry, L J & Wetzel, R (1987) *Protein Eng* **1**, 101–105
45 Rechsteiner, M , Rogers, S & Rote, K (1987) *Trends Biochem Sci* **12**, 390–394
46 Whitt, G S (1987) in *Isozymes Current Topics in Biological and Medical Research*, eds Rattazzi, M C , Scandalios, J G & Whitt, G S (Liss, New York), Vol 15, pp 1–26
47 Markert, C L (1984) *Cell Biochem Function* **2**, 131–134
48 Dabrowska, A & Gutowicz, J (1986) *Biochim Biophys Acta* **855**, 99–104
49 Gosti, F , Marty, M -C , Courvalin, J C , Maunoury, R & Bornens, M (1987) *Proc Natl Acad Sci USA* **84**, 1000–1004

CHAPTER 7

λ-CRYSTALLIN,

A MAJOR RABBIT LENS PROTEIN,

IS RELATED TO HYDROXYACYL-COENZYME A DEHYDROGENASES

John W.M. Mulders, Wiljan Hendriks,

W. Matthijs Blankesteijn, Hans Bloemendal & Wilfried W. de Jong

J. Biol. Chem. 263 (1988) 15462-15466.

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol 263 No 30 Issue of October 25 pp 15462-15466 1988

λ-Crystallin, a Major Rabbit Lens Protein, Is Related to Hydroxyacyl-coenzyme A Dehydrogenases*

(Received for publication June 1 1988)

John W. M. Mulders, Wiljan Hendriks, W. Matthijs Blankesteijn, Hans Bloemendal, and Wilfried W. de Jong

From the Department of Biochemistry University of Nijmegen, 6500 HB Nijmegen The Netherlands

It has recently been discovered that several lens proteins in birds and lower vertebrates are active enzymes or enzyme-related proteins (Wistow, G., Mulders, J. W. M., and de Jong, W. W. (1987) *Nature* 326, 622–624; Wistow, G., and Piatigorsky, J. (1987) *Science* 236, 1554–1556). We report here a novel lens protein, designated as λ-crystallin, that occurs in rabbit and hare. It constitutes 7–8% of the total lens protein and has a subunit molecular mass of 35 kDa. Sequencing of cDNA clones encoding rabbit λ-crystallin revealed 30% homology (at the amino acid sequence level) with L-3-hydroxyacyl-CoA dehydrogenase from pig mitochondria and 26% homology with enoyl-CoA hydratase-3-hydroxyacyl-CoA dehydrogenase from rat peroxisomes. Also, the presence of a putative β-α-β nucleotide-binding fold and low levels of non-lens expression are indicative of some enzymatic function for λ-crystallin (or highly related sequences) in non-lens tissues. λ-Crystallin thus represents the first example of an enzyme-related crystallin in lenses from mammalian species. The recruitment of enzymes as lens structural proteins apparently is an evolutionary strategy which has been applied independently in different lineages.

The vertebrate eye lens is a highly specialized tissue which is characterized by the presence of large amounts (up to 60% of the wet weight) of only a few, well-characterized, water-soluble proteins the crystallins (for reviews, see Bloemendal, 1981, Harding and Crabbe, 1984, Maisel, 1985, Wistow and Piatigorsky, 1988) The most ubiquitous lens proteins are the α-crystallins, which are related to the small heat shock proteins (Ingolia and Craig, 1982, de Jong *et al*, 1988) and to the major *Schistosoma* egg antigen (Nene *et al*, 1986), and the β/γ-crystallins, which show structural similarity to protein S from *Myxococcus xanthus* (Wistow *et al*, 1985) Recently, it has been shown that some major lens proteins with restricted occurrence in birds, reptiles, and other lower vertebrates are identical or similar to certain enzymes The first example of an enzyme which is expressed as a lens protein was ε-crystallin (Wistow *et al*, 1987) ε-Crystallin is present in large amounts (up to 23% of the total lens protein) in some birds and reptiles and has been shown to be authentic lactate dehydrogenase B4 (Wistow *et al*, 1987, Hendriks *et al*, 1988a) Other cases of vertebrate crystallins exhibiting sequence similarity to enzymes are δ and τ-crystallins, which share homology with argininosuccinate lyase (Wistow and Piatigorsky, 1987, Piatigorsky *et al*, 1988) and α enolase (Wistow and Piatigorsky, 1987), respectively, whereas ρ crystallin is related to aldose and aldehyde reductase (Carper *et al*, 1987) as well as to prostaglandin F synthase (Watanabe *et al*, 1988) These surprising findings evoke some major questions concerning the function, expression, and evolution of lens proteins

Up until now, no cases of enzyme related lens crystallins in mammals have been reported It appeared in fact that all mammalian lenses have comparable crystallin compositions, consisting only of α, β-, and γ crystallins (Bloemendal, 1981) Recently, however, a unique 36 kDa protein, called ζ crystallin, was isolated from guinea pig lenses (Huang *et al*, 1987), but its possible relation with any other protein has not yet been established In this paper, we characterize in detail a 35-kDa protein from rabbit lens It is designated as λ-crystallin because it has only been found in lenses of rabbit and hare (family Leporidae of the order Lagomorpha)

EXPERIMENTAL PROCEDURES

Materials—Rabbit and hare (*Lepus europaeus*) eyes were obtained on ice from a local butchery Pika (*Ochotona princeps*) lenses were sent in a saturated solution of guanidine HCl (de Jong *et al* 1984) by Dr M L Johnson (Seattle Washington)

Purification and Characterization of λ Crystallin—The water soluble proteins from rabbit lenses were isolated and separated on an Ultrogel AcA34 column (Pharmacia LKB Biotechnology Inc) using established procedures (van der Ouderaa *et al* 1973) Fractions containing λ crystallin were pooled and after carboxymethylation further purified on a Sephacryl S 200 column (Pharmacia LKB Biotechnology Inc) eluted with 10% acetic acid containing 6 M urea (Bindels *et al*, 1979) λ Crystallin was digested with trypsin and the obtained peptides were separated and isolated as described earlier (van der Ouderaa *et al* 1973) Amino acid analysis was performed on an Alpha Plus amino acid analyzer (Pharmacia LKB Biotechnology Inc) Amino acid sequences were determined on an Applied Biosystems Model 470A Protein Sequencer, equipped on line with a Model 120A phenylthiohydantoin analyzer

Preparation of Antibodies—Polyclonal antibodies against λ crystallin were prepared using an immunization scheme as described by Mulders *et al* (1988), except that only one booster injection was given In order to eliminate minor impurities, λ crystallin, purified as described above, was first electrophoresed on an SDS[1] polyacrylamide gel (Laemmli, 1970) and then eluted from the gel using isotachophoresis (Oferstedt *et al*, 1983) Antibody production was monitored by Western blotting as described by Mulders *et al* (1988)

Construction and Screening of Rabbit Lens cDNA Library and Sequence Analysis and Comparison—Total rabbit lens RNA was isolated following the procedure of Auffray and Rougeon (1980) Poly(A)$^+$ selection on oligo(dT) cellulose was done as described by Maniatis *et al* (1982) A cDNA expression library in λgt11 was prepared as described by Gubler and Hoffman (1983) Approximately

* This work was supported by the Netherlands Foundation for Chemical Research (Stichting Scheikundig Onderzoek Nederland) with financial aid from the Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelyk) The costs of publication of this article were defrayed in part by the payment of page charges This article must therefore be hereby marked "*advertisement*" in accordance with 18 U S C Section 1734 solely to indicate this fact

[1] The abbreviation used is SDS, sodium dodecyl sulfate

2×10^5 plaques were obtained upon infection of *Escherichia coli* Y1090. The method of Huynh *et al.* (1985) was used to screen the library with λ-crystallin antiserum for recombinants expressing λ-crystallin coding sequences. ^{125}I-Labeled donkey anti-rabbit immunoglobulins (16 μCi/μg; Amersham Corp.) were used as second antibody; autoradiography was done with intensifying screens at −70 °C. In order to obtain larger cDNA clones, the library was subsequently rescreened using ^{32}P-labeled λ-crystallin probes. Approximately 3% of the clones contained λ-crystallin-specific inserts. Nucleotide sequences were determined using the shotgun strategy of Lin *et al.* (1985) and the dideoxy sequencing method (Sanger *et al.*, 1980). Gel readings were recorded, compared, edited, and assembled by the programs of Staden (1977). Homology searches were done using the Swiss Prot data base (release 5.0).

RESULTS

Purification and Properties of a Major, 35-kDa Rabbit Lens Protein—Analysis of the water-soluble rabbit lens proteins reveals, in addition to the expected α-, β-, and γ-crystallins, a pronounced 35-kDa component which constitutes 7–8% of the total protein (Fig. 1). Considering the reported size heterogeneity of β-crystallin subunits (Berbers *et al.*, 1984), it was initially thought to be an exceptionally long β-chain. Upon gel filtration, the 35-kDa protein elutes between the β^H high and β^L low peaks (*cf.* Zigler and Sidbury, 1976), from which the molecular mass of the native protein was estimated to be 70 kDa. Therefore, the native protein seemed to be a dimer. However, using high performance gel permeation chromatography on Zorbax GF-450 and G-250, a calculated molecular weight of 105,000, between dimers and tetramers, was found.

In order to purify the protein, Ultrogel AcA34 column fractions containing the 35-kDa protein were pooled and chromatographed under denaturing conditions on a Sephacryl S-200 gel filtration column. Carboxymethylation prior to chromatography was found to result in a better separation between the 35-kDa protein and the low molecular mass β-crystallin subunits. Part of the carboxymethylated 35-kDa protein was used for preparative SDS-polyacrylamide gel electrophoresis in order to further purify it for antibody production. The obtained antiserum exclusively reacted with the 35-kDa protein in the water-soluble fraction of rabbit lens extract (*cf.* Fig. 5). The 35-kDa protein must therefore be immunologically distinct from α-, β-, and γ-crystallins. Also, using anti-β-, anti-ε, and anti-ζ-crystallin antisera, no reaction with the 35-kDa protein was detectable (data not shown).

In order to corroborate further that the 35-kDa protein is not related to other crystallins and to support the cDNA sequencing, tryptic fragments were isolated using high-voltage paper electrophoresis and chromatography. Sequences of two tryptic peptides were determined (*boxed* in Fig. 2) and compared with the Swiss Prot data base. No significant similarity with any known protein sequence, including those of the crystallins, was detected. We therefore decided that it was a novel protein and designated it as λ-crystallin.

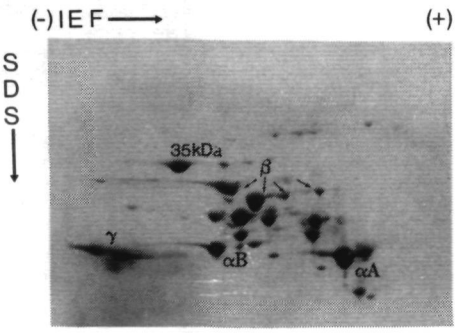


FIG. 1. **Two-dimensional gel electrophoresis (O'Farrell, 1975) of rabbit water-soluble lens proteins.** The spot marked *35 kDa* represents λ-crystallin. For isoelectric focusing (*IEF*), Ampholine (Pharmacia LKB Biotechnology, Inc.) at pH 3.5–10, 6–8, and 7–9 was mixed in a ratio of 10:3:3, respectively.

cDNA Cloning of λ-Crystallin—In order to obtain more information about the primary structure of λ-crystallin and to detect any possible relationship with other proteins, we used the anti-λ-crystallin serum to screen a λgt11 expression library prepared from rabbit lens mRNA. Several positive recombinants were obtained, and the complete nucleotide sequence of the longest cDNA insert (λRbLλ-4, 1424 base pairs) was determined (Fig. 2). It contains a poly(A) addition signal at position 1404, an initiation consensus sequence at position 61, and a stop codon at position 1021, generating an open reading frame of 957 nucleotides, encoding a protein of 319 amino acids (excluding the initiation methionine) with a molecular weight of 35,101. In addition to the two sequenced peptides (*boxed* in Fig. 2), the deduced amino acid compositions of the *underlined* peptides were also found to match exactly the amino acid compositions of purified tryptic peptides of λ-crystallin. Both the deduced N- and C-terminal peptides were recognized among the tryptic peptides recovered from the protein, indicating that no post-translational removal of signal sequences occurs in λ-crystallin. The methionine corresponding to the initiation codon was not found in the N-terminal peptide. According to the deduced sequence, the N-terminal tryptic peptide should be uncharged at pH 6.5. It actually has, however, a charge of −1, probably like most crystallins by blocking of the α-amino group by N^{α}-acetylation. N-terminal alanine is indeed commonly N^{α}-acetylated (Driessen *et al.*, 1985).

λ-Crystallin Is Homologous to L-3-Hydroxyacyl-CoA Dehydrogenase and Enoyl-CoA Hydratase-3-Hydroxyacyl-CoA Dehydrogenase—The deduced complete amino acid sequence was again compared with the Swiss Prot protein sequence data base. Similarity was found between the λ-crystallin sequence and porcine L-3-hydroxyacyl-CoA dehydrogenase, a mitochondrial enzyme involved in fatty acid metabolism (Bitar *et al.*, 1980) of which the tertiary structure was recently established (Birktoft *et al.*, 1987). A lesser similarity exists with rat enoyl-CoA hydratase-3-hydroxyacyl-CoA dehydrogenase, a bifunctional peroxisomal enzyme also involved in fatty acid metabolism (Osumi *et al.*, 1985). Although the homology is low (Fig. 3; 30 and 26%, respectively), scrambling tests (performed by Dr. R. F. Doolittle, University of California, San Diego) showed that the alignment is significant. Hydroxyacyl CoA-dehydrogenase is dimeric and has, like λ-crystallin, a subunit molecular weight of 35,000, whereas the peroxisomal bifunctional enzyme has a subunit molecular weight of 78,500.

Since distant (but distinct) homology was found with these two CoA dehydrogenases, we looked for the presence of the consensus sequence of a β-α-β fold, which has been shown to bind the ADP moiety of NAD or FAD (Wierenga *et al.*, 1986). In this "fingerprint," 3 glycines and 1 acidic residue at the C-terminal end of the consensus are absolutely required. Some variation, however, at the first position and at the position of the hydrophobic residues is allowed. λ-Crystallin indeed contains the β-α-β fold fingerprint (Fig. 3): 9 out of 11 residues, including the obligatory glycines, at positions 12, 14, and 17 and aspartic acid at position 35 match the consensus sequence. An aspartic acid residue at the first position of the β-α-β fold

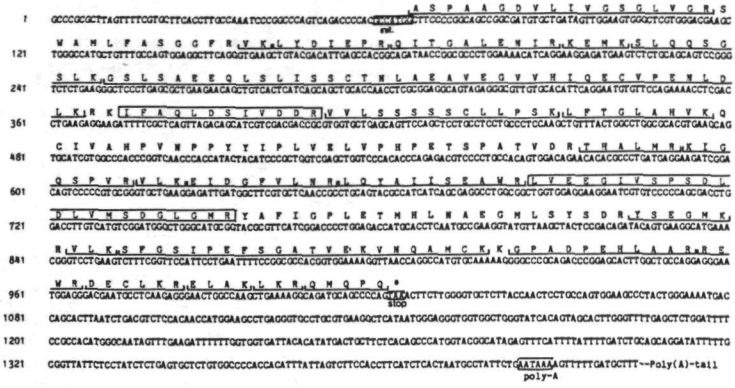


FIG. 2. **Complete nucleotide sequence of cDNA insert from clone λRbLλ-4 and deduced amino acid sequence of rabbit λ-crystallin.** The initiation ATG codon and its consensus (*init.*), the stop codon, and the putative poly(A) addition signal are indicated. *Boxed* amino acid sequences are tryptic peptides that were sequenced. Tryptic peptides of which the amino acid compositions were determined are *underlined*. No methionine was found in the N-terminal peptide.

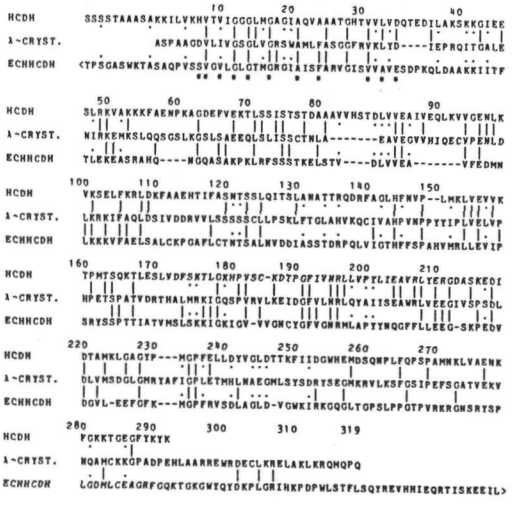


FIG. 3. **Comparison of amino acid sequences of pig mitochondrial L-3-hydroxyacyl-CoA dehydrogenase (*HCDH*), rabbit λ-crystallin (λ-*CRYST.*) and rat peroxisomal enoyl-CoA hydratase-3-hydroxyacyl-CoA dehydrogenase (*ECHH-CDH*).** The numbering of amino acid residues is given for λ-crystallin; amino acids 282–619 of the enoyl-CoA hydratase-3-hydroxyacyl-CoA dehydrogenase chain are shown. The *asterisks* indicate those residues which may form a β-α-β fold for nucleotide binding. *Vertical lines* indicates homologous positions between λ-crystallin and either enzyme. *Black dots* indicate amino acids that are identical in both enzymes, but not in λ-crystallin. Homology is calculated for positions 1–246 and is 30 and 26% for λ-crystallin *versus* L-3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase-3-hydroxyacyl-CoA dehydrogenase, respectively, and 32% between the two enzymes.

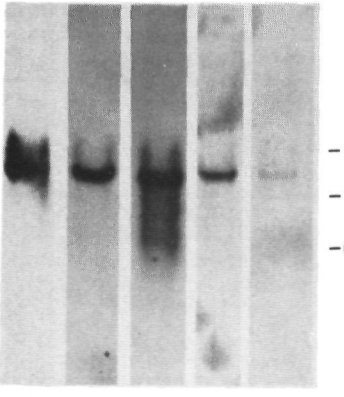


FIG. 4. **λ-Crystallin expression in non-lens rabbit tissues.** RNA was isolated from rabbit tissues as described (Auffray and Rougeon, 1980). Total RNA fractions from liver, spleen, heart, and kidney or poly(A)$^+$ selected lens RNA was glyoxylated, fractionated on 1% agarose gels, and transferred to Hybond N (Amersham Corp.) filters. As a probe, the cDNA insert from clone λRbLλ-1 (743 base pairs, nucleotides 77–820) was used. Hybridization of Northern blots was as described by Church and Gilbert (1984) with the addition of 100 μg of single-stranded herring sperm DNA/ml of hybridization mixture. Blots were washed twice for 30 min at 65 °C with 0.5 M Na_2HPO_4, pH 7.2, containing 1 mM EDTA and 1% SDS. *Lane 1*, lens, 2 μg of poly(A)$^+$ RNA; *lane 2*, liver, 15 μg of total RNA; *lane 3*, kidney, 15 μg of total RNA; *lane 4*, spleen, 15 μg of total RNA; *lane 5*, heart, 15 μg of total RNA. In all cases, except in the kidney RNA, where some degradation had occurred, a single 1.65-kilobase transcript was identified. Exposure times were 2 h (*lane 1*), 18 h (*lanes 2* and *3*) and 3 days (*lanes 4* and *5*).

has also been reported for glutathione reductase (Schulz *et al.*, 1982), whereas phenylalanine at position 24 is also found in enoyl-CoA hydratase-3-hydroxyacyl-CoA dehydrogenase (Osumi *et al.*, 1985) without disturbing NAD binding. Taken together, these data indicate that λ-crystallin may contain an NAD- or FAD-binding structure. In order to corroborate this finding, 5′-AMP-Sepharose chromatography of total rabbit lens water-soluble proteins was performed. λ-Crystallin was, however, exclusively found in the unbound fraction (data not shown), indicating that, under these conditions, no effective NAD binding seems to occur.

λ-Crystallin Is Not Lens-specific—Because λ-crystallin shows homology to hydroxyacyl-CoA dehydrogenases and probably contains an NAD-binding fold, it may well be an enzyme itself. We therefore analyzed the occurrence of λ-crystallin or λ-crystallin-like mRNA and proteins in non-lens tissues of rabbit by Northern and Western blotting. In Fig. 4, a Northern blot, hybridized with the insert from clone λRbLλ-1, clearly shows that RNAs from lens, liver, kidney, spleen, and heart all contain a single hybridizing transcript of 1.65 kilobases. As compared with lens, the level of expression is considerably lower in liver and kidney and much lower in heart and spleen. The size of the identified transcript implies

that λRbLλ-4 could represent a full-length cDNA clone, assuming an average length of about 200 bases for the poly(A) tail. Also, using Western blotting with the anti-λ-crystallin serum, small amounts of an approximately 35–37-kDa protein are detectable in heart (Fig. 5*C*), skeletal muscle, spleen, and liver extracts (not shown).

λ-Crystallin Is Found in Rabbit and Hare—In order to study the distribution of λ-crystallin as an abundant lens protein, Western blots containing the water-soluble lens proteins from different vertebrate species were incubated with the anti-λ-crystallin serum. Fig. 5*A* shows the Coomassie Blue-stained electrophoretic profile of rabbit, hare, and pika lens extracts. Rabbits and hares belong to the family Leporidae of the order Lagomorpha, whereas pikas belong to the only other family, Ochotonidae, of this order. On the corresponding Western blot (Fig. 5*B*), the presence of λ-crystallin in both rabbit and hare is obvious, whereas no λ-crystallin is detectable in pika. When testing lens extracts from other species with prominent approximately 35-kDa proteins, such as guinea pig, the rodent *Ctenodactylus gundi* (Hendriks *et al.*, 1988b), duck, and frog, none of them showed any reaction with the anti-λ-crystallin serum. No other mammalian species with a conspicuous 35-kDa lens protein have been encountered in our previous reports (de Jong *et al.*, 1984; Hendriks *et al.*, 1988b).

Genomic Organization of λ-Crystallin Sequences—The genomic complexity of λ-crystallin sequences and the occurrence in other mammalian species were analyzed using Southern blots of *Pst*I, *Eco*RI, *Hin*dIII, or *Bam*HI digests of rabbit, human, and bovine chromosomal DNAs. A rather complex hybridization pattern was obtained for all digests using the λRbLλ-4 insert as a probe (data not shown). It was therefore concluded that λ-crystallin-like genomic sequences are present in all three species, whereas high expression in the lens is only found in rabbit. From the hybridization patterns, it was not possible to decide whether λ-crystallin is encoded by a single copy gene or whether multiple λ-crystalline-like-sequences are present in the mammalian genome.

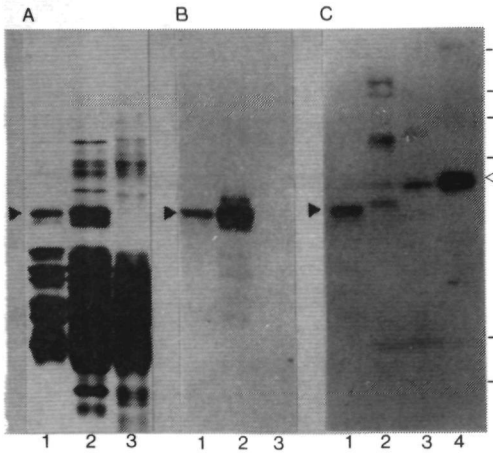


FIG. 5. **λ-Crystallin is found only in rabbit and hare and occurs outside the lens.** *A*, Coomassie Blue-stained SDS gel of rabbit (*lane 1*), hare (*lane 2*), and pika (*lane 3*) water-soluble lens proteins. *B*, immunoblot incubated with the antiserum against λ-crystallin. Samples are the same as described for *A*. *Solid arrowheads* indicate λ-crystallin. λ-Crystallin is present only in rabbit and hare and is not detectable in pika. *C*, immunoblot of total soluble protein of rabbit lens (*lane 1*) and heart (*lane 2*) using the anti-λ-crystallin serum. Upon carboxymethylation of heart (*lane 3*) and lens (*lane 4*) extracts, λ-crystallin migrates at an apparent molecular mass of 42 kDa (*open arrowhead*). The apparent molecular mass of the anti-λ-crystallin cross-reacting protein in heart tissue is slightly higher in comparison with λ-crystallin from lens. The identity of the other, higher molecular mass cross-reacting protein species in heart tissue is not known.

DISCUSSION

Enzyme-related crystallins had until now only been identified in the lenses of birds and lower vertebrates. In mammalian lenses, only the classical α-, β-, and γ-crystallins were apparently found. This may be a fundamental property of mammalian lenses, but it can also be a biased impression resulting from limited and inadequate comparative studies of mammalian lens proteins. It is indeed possible that in certain mammalian species inconspicuous additional lens proteins are hidden among the common α-, β-, and γ-crystallins, like ζ-crystallin in guinea pig (Huang *et al.*, 1987).

Another example is a 33-kDa rabbit lens protein, which elutes on a gel filtration column as a shoulder in front of the β low-peak, and was originally described as a β-crystallin, capable of forming homoaggregates (Zigler and Sidbury, 1976). In one study, a tetrameric quaternary structure was assumed (Mostapafour and Reddy, 1978), whereas in an other study, an $s_{20,w}$ of 3.7 S (suggesting a dimer) was reported (Liem-The and Hoenders, 1974). In our experiments using gel filtration over Ultrogel AcA34, this protein, now characterized as λ-crystallin, elutes just before the β low-peak, indicating a molecular weight of approximately 70,000, as expected for a dimeric protein, whereas high performance gel permeation indicates a dimeric or tetrameric protein. Both dimeric and tetrameric forms of, for instance, glyceraldehyde-3-phosphate dehydrogenase have been reported (Grosse *et al.*, 1986), and α-enolase has been isolated in monomeric as well as dimeric forms (Giallongo *et al.*, 1986).

As for other enzyme-related crystallins (Carper *et al.*, 1987; Wistow *et al.*, 1987; Wistow and Piatigorsky, 1987), λ-crystallin (or λ-crystallin-like) expression is also found in non-lens tissues. The finding that λ-crystallin contains a putative NAD-binding site, despite the absence of binding to 5′-AMP-Sepharose, further corroborates that λ-crystallin is closely related or even identical to some, as yet unidentified, active enzyme. Preliminary data[2] show that "λ-crystallin" from liver is, again like hydroxyacyl-CoA dehydrogenase, a mitochondrial protein. The loss of mitochondria during lens cell differentiation (Harding and Crabbe, 1984) precludes, however, the occurrence of λ-crystallin as a mitochondrial protein in the lens. If λ-crystallins from lens and liver are the same gene products, λ-crystallin contains, if at all, only a short signal sequence to be cleaved off after translocation into mitochondria because the putative β-α-β fold starts at residue 7 of the lens protein. It is also possible that λ-crystallin does not contain a cleavable presequence, as is also the case for the mitochondrial matrix proteins 2-isopropylmalate synthase (Hampsey *et al.*, 1983) and carbamoyl-phosphate synthetase (Mori *et al.*, 1979). The third possibility is that λ-crystallin from lens and λ-crystallin from non-lens tissues are highly homologous but distinct gene products. cDNA cloning and sequencing of λ-crystallin-related mRNA from other tissues will discriminate between these alternatives. The complexity of the hybridization patterns of genomic DNA with the λ-crystallin probe is compatible with the existence of multiple genes.

Major questions arising from this and other studies (Carper

[2] J. W. M. Mulders and W. W. de Jong, unpublished observations.

et al, 1987 Wistow *et al* 1987 Wistow and Piatigorsky 1987) concern the functional and evolutionary significance of the presence of λ crystallin in Leporidae and of other crystallin enzymes in lenses from other species How did an enzyme become a major lens protein in just a single evolutionary lineage, and what is the advantage if any of the presence of this particular protein in the lens? It may be a form of 'evolutionary tinkering (Jacob 1983) the creative use in evolution of existing DNA and protein sequences in new combinations and for novel functions One may assume that the expression of λ crystallin which probably is identical or closely related to a widespread housekeeping enzyme has dramatically increased in the common ancestral lineage of the family Leporidae after its divergence from the family Ochotonidae approximately 35 million years ago (McKenna 1982) This may have occurred by the chance acquisition of lens specific promoter or enhancer sequences λ Crystallin must already have had appropriate structural properties to be acceptable as a lens protein, namely stability and the possibility of longevity and close packing (Delaye and Tardieu 1983) It is indeed difficult to envisage that λ crystallin provided a specific selective, that is reproductive, advantage over the other crystallins that were already present It therefore seems more plausible that this change of expression and the maintenance of λ crystallin during the divergent evolution of the family Leporidae represent selectively neutral events

Acknowledgments—We thank Frik ten Kergert and Inge Gerads for their contributions Marlies Versteeg for amino acid analyses Dr Jack Schalken for his advice in preparing the cDNA library Dr J S Zigler for the ζ crystallin antiserum Dr M I Johnson for pika lenses and Dr R F Doolittle for performing a scrambling test The help of Dr Jack Leunissen as well as the use of the services and facilities of the Dutch Computer Assisted Organic Synthesis/Computer Assisted Macromolecular Modelling Center are gratefully acknowledged Amino acid sequences were determined at the SON gas phase sequenator facility (State University of Leiden)

REFERENCES

Auffray C and Rougeon F (1980) *Eur J Biochem* **107**, 303 314

Berbers G A M Hoekman W A Bloemendal H de Jong W W Kleinschmidt T and Braunitzer G (1984) *Eur J Biochem* **139**, 467 479

Bindels J G Siezen R J and Hoenders H J (1979) *Ophthalmic Res* **11** 441 452

Birktoft J J Holden H M Hamlin R Xuong N H and Banaszak L J (1987) *Proc Natl Acad Sci U S A* **84**, 8262 8266

Bitar K G Perez Aranda A and Bradshaw R A (1980) *FEBS Lett* **16**, 196 198

Bloemendal H (ed) (1981) *Molecular and Cellular Biology of the Eye Lens* John Wiley & Sons New York

Carper D Nishimura C Shinohara T Dietzchold B Wistow G Craft C Kador P and Kinoshita J H (1987) *FEBS Lett* **220**, 209 213

Church G M and Gilbert W (1984) *Proc Natl Acad Sci U S A* **81**, 1991 1995

de Jong W W, Zweers A Versteeg M and Nuy Terwindt E C (1984) *Eur J Biochem* **141**, 131-140

de Jong W W Leunissen J A M Leenen P J M Zweers A and Versteeg M (1988) *J Biol. Chem* **263**, 5141-5149

Delaye M and Tardieu A (1983) *Nature* **302**, 415-417

Driessen H P C de Jong W W Tesser G I and Bloemendal H (1985) *CRC Crit Rev Biochem* **18**, 281 325

Giallongo A Feo S Moore R Croce C M and Showe L C (1986) *Proc Natl Acad Sci U S A* **83**, 6741-6744

Grosse F Nasheuer H P Scholtissek S and Schomburg U (1986) *Eur J Biochem* **160** 459 467

Gubler U and Hoffman B J (1983) *Gene (Amst)* **25** 263 269

Hampsey D M Lewin A S and Kohlhaw G B (1983) *Proc Natl Acad Sci U S A* **80** 1270-1274

Harding J J and Crabbe M J C (1984) in *The Eye* (Davson H ed) Vol 1b 3rd Ed pp 207 492 Academic Press New York

Hendriks W Mulders J W M Bibby M A Slingsby C Bloemendal H and de Jong W W (1988a) *Proc Natl Acad Sci U S A in press*

Hendriks W Sanders J de Leij L Ramaekers F Bloemendal H and de Jong W W (1988b) *Eur J Biochem* **174** 133 137

Huang Q L Russell P Stone S H and Zigler J R Jr (1987) *Curr Eye Res* **6** 729 732

Huynh T V Young R A and Davis R W (1985) in *DNA Cloning Techniques A Practical Approach* (Glover D M ed) pp 49 78 IRL Press Oxford

Ingolia T D and Craig E (1982) *Proc Natl Acad Sci U S A* **79**, 2360 2364

Jacob F (1982) in *Evolution from Molecules to Men* (Bendall D S ed) pp 131 144 Cambridge University Press Cambridge

Laemmli U K (1970) *Nature* **227** 680 685

Liem The K N and Hoenders H J (1974) *Exp Eye Res* **18**, 143-152

Lin H C Lei S P and Wilcox G (1985) *Anal Biochem* **147**, 114 119

Maisel H (ed) (1985) *The Ocular Lens Structure Function and Pathology* Marcel Dekker Inc New York

Maniatis T Fritsch E F and Sambrook J (1982) *Molecular Cloning A Laboratory Manual* Cold Spring Harbor Laboratory Cold Spring Harbor NY

McKenna M C (1982) *Geobios Mem Spec* **6**, 213-223

Mori M Morris S M and Cohen P P (1979) *Proc Natl Acad Sci U S A* **76** 3179 3183

Mostafavour M K and Reddy V N (1978) *Invest Ophthalmol Visual Sci* **17** 660 666

Mulders J W M Voorter C F M Lamers C de Haard Hoekman W A Montecucco C van de Ven W J M Bloemendal H and de Jong W W (1988) *Curr Eye Res* **7** 207 219

Nene V Dunne D W Johnson K S Taylor D W and Cordingley J S (1986) *Mol Biochem Parasitol* **21** 179-188

O Farrell P H (1975) *J Biol Chem* **250** 4007-4021

Oferstedt I G Sundelin J and Johansson G (1983) *Anal Biochem* **134**, 361 367

Osumi T Ishii N Hijikata M Kamijo K Ozasa H Furuta S Miyazawa S Kondo K Inoue K Kagamiyama H and Hashimoto T (1985) *J Biol Chem* **260** 8905 8910

Piatigorsky J O Brien W F Norman B L Kalumuck K Wistow G J Borras T Nickerson J M and Wawrousek E F (1988) *Proc Natl Acad Sci U S A* **85** 3479 3483

Sanger F Coulson A R Barrell B G Smith A J H and Roe B A (1980) *J Mol Biol* **143**, 161-178

Schulz G E Schirmer R H and Pai E F (1982) *J Mol Biol* **160** 287 308

Staden R (1977) *Nucleic Acids Res* **4**, 4037-4051

van der Ouderaa F J de Jong W W and Bloemendal H (1973) *Eur J Biochem* **39**, 207 222

Watanabe K Fuju Y Nakayama K Ohkubo H Kuramitsu S Kagamiyama H Nakanishi S and Hayashi O (1988) *Proc Natl Acad Sci U S A* **85**, 11 15

Wierenga R K Terpstra P and Hol W G J (1986) *J Mol Biol* **187**, 101 107

Wistow G and Piatigorsky J (1987) *Science* **236**, 1554-1556

Wistow G and Piatigorsky J (1988) *Annu Rev Biochem* **57**, 479 504

Wistow G Summers L and Blundell T (1985) *Nature* **315**, 771 773

Wistow G Mulders J W M and de Jong W W (1987) *Nature* **326**, 622 624

Zigler J S Jr and Sidbury J B Jr (1976) *Comp Biochem Physiol B Comp Biochem* **53**, 349-355

SUMMARY

SAMENVATTING

SUMMARY

The abundant lens proteins, called crystallins, are evolutionarily well-conserved. This is thought to indicate that these proteins serve important functions in the lens. Unfortunately, there are no assays available to reveal and evaluate the functional properties of these lens structural proteins. However, the comparison of crystallins from different species can give clues about the structural and functional requirements imposed upon these proteins. We therefore focussed our attention on various aberrant representatives of the most common lens protein, α-crystallin, to gain insight in its structure and function.

α-Crystallin is a multimeric protein composed of subunits that are encoded by two single-copy genes: the αA- and the αB-crystallin gene. In chapter 2 the complete structure of the hamster αA-crystallin gene is presented. This gene does not only code for the normal primary gene product αA, but also gives rise to an elongated chain, αA^{ins}, by means of alternative splicing of its second exon. The first exon is internally duplicated, and the two last exons contain sequences homologous to the small heat-shock proteins. A tentative model for the evolution of the αA-crystallin gene is presented, but additional sequences of α-crystallin subunits and small heat-shock proteins have now permitted a more precise reconstruction (chapter 1, ref. 34).

To determine the functional constraints that have shaped the αA-crystallin gene product during vertebrate evolution, the αA-crystallin sequence of the blind mole rat has been deduced. This rodent has degenerated eyes as an adaptation to a subterranean way of life for more than 25 million years. In chapter 3 it is argued that the nonsynonymous substitution rate of the αA-crystallin gene in the mole rat lineage is not accelerated to the extent that might be expected if the gene were now without any function. It is suggested that the remaining evolutionary constraints may reflect the need for α-crystallin expression as an indispensable component in the developmental program of the eye. The recent observation on the extra-lenticular occurrence of αB-crystallin (Bhat, S.P. et al. [1988] Inv. Ophth. Vis. Sci. 29, p. 426) of course strengthens another possibility, namely that also αA-crystallin fulfils some as yet unknown role outside the lens.

In chapter 4 a study is described in which monoclonal antibodies are used to examine the occurrence in the lenses of different vertebrates of the elongated αA^{ins} chain. It appears that the mutational event resulting in the

alternative splicing pattern of the αA-crystallin gene, took place at least 70 million years ago. This phenomenon has persisted in rodents and some other, unrelated mammals, but disappeared again in most mammalian lineages. The influence of the 23-residue insertion peptide on the structure of the αA^{ins} subunit and the α-crystallin aggregate, is investigated in chapter 5. The data suggest that αA, αA^{ins}, and αB occupy equivalent sites in the aggregate. Moreover, iodination of surface residues revealed no differences between αA and αA^{ins}. It seems that the insertion is nothing but a connecting peptide between the N-terminal and C-terminal domain of the αA-crystallin chain.

Recently some crystallins were identified as being identical with or related to certain enzymes. This, of course, raises the question what makes these proteins suited for functioning in the lens. It also provides a possibility to study these lens proteins by monitoring their enzymatic activity. Chapter 6 deals with the question whether or not duck lens ε-crystallin and duck heart lactate dehydrogenase (LDH) B4 are the product of the same gene. By using cDNA clones of duck ε-crystallin it is demonstrated that the single-copy Ldh-B gene in duck is overexpressed in lens tissue. This duck lens LDH-B4 is subsequently posttranslationally modified, explaining the observed minor differences between LDH-B4 and ε-crystallin. Possible adaptations to the additional function in the lens are discussed. Finally, in chapter 7 the first example of an enzyme-related crystallin in mammals is described. A 35 kDa protein that is abundant in rabbit and hare lenses, called λ-crystallin, was studied. Sequence analysis of rabbit λ-crystallin cDNA clones reveals significant homology with hydroxyacyl CoA dehydrogenases. Low levels of λ-crystallin or λ-crystallin-like mRNA and protein are detectable in other tissues, suggesting an enzymatic role outside the lens.

SAMENVATTING

De crystallines, eiwitten die in zeer grote hoeveelheden in de ooglens voorkomen, zijn evolutionair gezien zeer conservatief. Dit betekent waarschijnlijk dat deze eiwitten belangrijke functies vervullen in de lens. Helaas zijn er geen tests voorhanden om de functionele eigenschappen van de meeste structurele lenseiwitten te bepalen. Het vergelijken van crystallines van verschillende dieren kan echter indirecte aanwijzingen geven over de structurele en functionele voorwaarden waaraan deze eiwitten moeten voldoen. Wij hebben ons daarom gericht op verschillende afwijkende vertegenwoordigers van α-crystalline, een lenseiwit dat bij vrijwel alle vertebraten wordt gevonden, teneinde inzicht te verkrijgen in de structuur en functie van dit eiwit.

α-Crystalline is een multimeer eiwit dat bestaat uit polypeptideketens die gecodeerd worden door twee, slechts éénmaal in het haploid genoom voorkomende, genen: het αA- en het αB-crystalline-gen. In hoofdstuk 2 wordt de structuur van het hamster αA-crystalline-gen gepresenteerd. Dit gen codeert niet alleen voor αA, het normale primaire genproduct, maar leidt ook tot een langere keten, αA^{ins}-crystalline, door alternatieve "splicing" van het tweede exon. Het eerste exon is intern gedupliceerd, en de laatste twee exonen bevatten sequenties die homologie vertonen met de kleine "heat-shock"-eiwitten. Een voorlopig model van de evolutie van het αA-crystalline-gen wordt gepresenteerd, maar additionele sequenties van α-crystalline-ketens en kleine "heat-shock"-eiwitten laten inmiddels een meer precieze reconstructie toe (hoofdstuk 1, referentie 34).

Om de functie-eisen te bepalen die het αA-crystalline-genproduct hebben gevormd gedurende de evolutie van de vertebraten, is de αA-crystallinesequentie van de blinde molrat opgehelderd. Dit knaagdier heeft gedegenereerde ogen als gevolg van een minstens 25 miljoen jaar-lange aanpassing aan een ondergrondse leefwijze. In hoofdstuk 3 wordt betoogd dat de mutatiesnelheid van het αA-crystalline in het molratgeslacht niet zo dramatisch is versneld als mag worden verwacht indien het eiwit geen functie meer heeft. Gesuggereerd wordt dat de overgebleven evolutionaire druk de behoefte weerspiegelt aan α-crystalline-expressie als onmisbaar onderdeel in de

ontwikkeling van het oog. De recente bevinding dat αB-crystalline ook in weefsels buiten de lens aantoonbaar is (Bhat, S.P. et al. [1988] Inv. Ophth. Vis. Sci. 29, p. 426) sterkt natuurlijk een andere mogelijkheid, namelijk dat ook αA-crystalline een nog onbekende rol buiten de lens vervult.

In hoofdstuk 4 wordt een studie beschreven waarin monoclonale antilichamen worden gebruikt om te bepalen bij welke vertebraten de verlengde αA^{ins}-keten in de lens voorkomt. Het blijkt dat de mutatie die resulteerde in het alternatieve splicingspatroon van het αA-crystalline-gentranscript, minstens 70 miljoen jaar geleden plaatsvond. Dit fenomeen bleef gehandhaafd in knaagdieren en enkele andere zoogdieren, maar verdween weer in de meeste zoogdierfamilies. De invloed van de 23 aminozuren-lange insertie op de structuur van de αA^{ins}-crystalline-keten en van het α-crystalline-aggregaat is onderzocht in hoofdstuk 5. De resultaten suggereren dat αA, αA^{ins} en αB gelijkwaardige posities innemen in het aggregaat. Verder laat jodering van oppervlakteresiduen geen verschillen zien tussen αA en αA^{ins}. De insertie lijkt niets meer te zijn dan een verbindend peptide tussen het aminoterminale en carboxyterminale domein van de αA-crystalline-keten.

Recent is gebleken dat enkele crystallines verwant zijn met of zelfs identiek zijn aan bepaalde enzymen. Dit roept natuurlijk de vraag op wat deze eiwitten geschikt maakt om te functioneren in de lens. Het geeft ook een mogelijkheid om lenseiwitten te bestuderen aan de hand van hun enzymatische activiteit. Hoofdstuk 6 behandelt de vraag of ε-crystalline uit de eendelens en lactaat-dehydrogenase (LDH) B4 uit eendehart producten zijn van een en hetzelfde gen. Gebruik makend van cDNA-clonen van eend ε-crystalline kon aangetoond worden dat een uniek Ldh-B-gen in het genoom van de eend zeer sterk tot expressie komt in lensweefsel. Dit eendelens-LDH B4 wordt vervolgens posttranslationeel gemodificeerd, hetgeen de kleine verschillen verklaart tussen ε-crystalline en LDH B4. Mogelijke aanpassingen aan de nieuwe functie in de lens worden besproken. Tenslotte wordt in hoofdstuk 7 het eerste voorbeeld van een enzym-gerelateerd crystalline bij zoogdieren gerapporteerd. Een 35 kDa-eiwit dat in grote hoeveelheden voorkomt in konijne- en haze-lenzen werd λ-crystalline genoemd. De cDNA-sequentie van konijn λ-crystalline vertoonde een significante homologie met hydroxyacyl-CoA-dehydrogenases. Kleine hoeveelheden λ-crystalline of λ-crystalline-achtige mRNA's en eiwitten zijn aantoonbaar in andere weefsels dan de lens, hetgeen een enzymatische rol buiten de lens suggereert.

LIST OF PUBLICATIONS

- Quax, W., Vree Egberts, W., Hendriks, W., Quax-Jeuken, Y. & Bloemendal, H. (1983) The structure of the vimentin gene. Cell 35, 215-223.

- van den Heuvel, R., Hendriks, W., Quax, W. & Bloemendal, H. (1985) Complete structure of the hamster αA crystallin gene: Reflection of an evolutionary history by means of exon shuffling. J. Mol. Biol. 185, 273-284.

- Bloemendal, H., van den Heuvel, R., Quax-Jeuken, Y., Hendriks, W., van Rens, G. & Quax, W. (1986) The two genes of α-crystallin: Conservation of exon-intron distribution, nonlinkage and homology with heat-shock protein genes. In: The Lens: Transparency and Cataract (Duncan, G., ed.) Topics in aging research in Europe (Eurage, Rijswijk) Vol. 6, pp. 165-169.

- de Jong, W.W. & Hendriks, W. (1986) The eye lens crystallins: Ambiguity as evolutionary strategy. J. Mol. Evol. 24, 121-129.

- Hendriks, W., Leunissen, J.A.M., Nevo, E., Bloemendal, H. & de Jong, W.W. (1987) The lens protein αA-crystallin of the blind mole rat, Spalax ehrenbergi: Evolutionary change and functional constraints. Proc. Natl. Acad. Sci. USA 84, 5320-5324.

- Hendriks, W., Bloemendal, H. & de Jong, W.W. (1987) Alu B2 type repeats in the αA-crystallin locus in the mole rat (Spalax ehrenbergi) genome. Nucl. Acids Res. 15, 9093.

- Hendriks, W., Sanders, J., de Leij, L., Ramaekers, F., Bloemendal, H. & de Jong, W.W. (1988) Monoclonal antibodies reveal evolutionary conservation of alternative splicing of the αA-crystallin primary transcript. Eur. J. Biochem. 174, 133-137.

- Hendriks, W., Mulders, J.W.M., Bibby, M.A., Slingsby, C., Bloemendal H. & de Jong, W.W. (1988) Duck lens ε-crystallin and lactate dehydrogenase B4 are identical: A single-copy gene product with two distinct functions. Proc. Natl. Acad. Sci. USA 85, 7114-7118.

- Mulders, J.W.M., Hendriks, W., Blankesteijn, W.M., Bloemendal, H. & de Jong, W.W. (1988) λ-crystallin, a major rabbit lens protein, is related to hydroxyacyl CoA dehydrogenases. J. Biol. Chem. 263, 15462-15466.

- de Jong, W.W., Leunissen, J.A.M., Hendriks, W. & Bloemendal, H. (1988) Molecular evolution of α-crystallin: in quest of a function. In: Molecular Biology of the Eye: Genes, Vision & Ocular Disease (Piatigorsky, J., Zelenka, P. & Shinohara, T., eds.) UCLA Symposia on Molecular and Cellular Biology, new series (Liss, New York) vol. 88, in press.

- Hendriks, W., Weetink, H., Voorter, C.E.M., Sanders, J., Bloemendal, H. & de Jong, W.W. (1988) Structural equivalence of subunits in the rat α-crystallin aggregate. Submitted for publication.

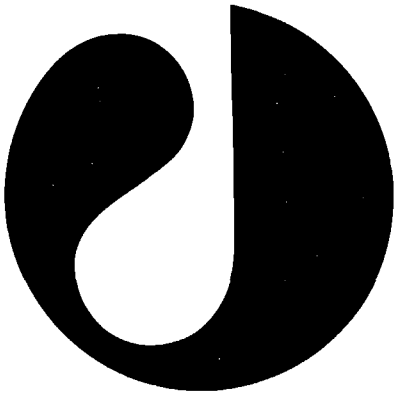
CURRICULUM VITAE

Wiljan Hendriks werd geboren op 7 oktober 1959 te Ruigoord, een klein plaatsje in de gemeente Haarlemmerliede en Spaarnwoude, even ten westen van Amsterdam. Na het behalen van het VWO-diploma aan het Thomas à Kempis College te Arnhem werd in augustus 1977 begonnen met de studie scheikunde aan de Katholieke Universiteit te Nijmegen. Het kandidaats examen werd in september 1980 afgelegd. In juni 1984 werd het doctoraal examen afgelegd met als uitgebreid bijvak Biofysische Chemie (Drs. H. Janssen, Prof. Dr. C.W. Hilbers) en als hoofdvak Biochemie (Dr. W. Quax, Prof. Dr. H. Bloemendal).

Vanaf juli 1984 tot juli 1988 was hij in dienst van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) (Stichting S.O.N., Werkgemeenschap Eiwitten), en werkzaam op het Laboratorium voor Biochemie van de Katholieke Universiteit te Nijmegen. In deze periode werd, onder leiding van Dr. W.W. de Jong en Prof. Dr. H. Bloemendal, het in dit proefschrift beschreven onderzoek verricht. Tevens werd een bijdrage geleverd aan het praktisch biochemie-onderwijs aan derde jaars studenten scheikunde.

Dankzij reisbeurzen van F.E.B.S. (Federation of European Biochemical Societies), NWO en SON werden de FEBS Advanced Course "Genomic Organization and Evolution" (Kreta, september 1986), de KNAW workshop "Recent Developments in Molecular Evolution" (Amsterdam, juni 1987), het UCLA Symposium "Molecular Biology of the Eye; Genes, Vision and Ocular Diseases" (Santa Fe, februari 1988), de Arolla Workshop "Mechanisms of Gene Regulation and Development" (Arolla, augustus 1988) en de FEBS Workshop "Evolutionary Tinkering in Gene Expression" (Spetses, september 1988) bijgewoond. Werkbezoeken, mogelijk gemaakt door financiële steun van SON en EURAGE, werden gebracht aan de afdeling Immuungenetica van het Max Planck Instituut te Tübingen (Prof. Dr. J. Klein, Dr. F. Figueroa, oktober 1985) en de afdeling Kristallografie van het Londense Birkbeck College (Prof. Dr. T.L. Blundell, Dr. C. Slingsby, december 1987).

Vanaf juli 1988 is hij werkzaam als wetenschappelijk medewerker in de werkgroep van Dr. W.W. de Jong.



Dit proefschrift werd mede mogelijk gemaakt door

PHARMACIA NEDERLAND B.V.

Pharmacia Nederland B.V.
Pharmacia LKB Biotechnology
Postbus 17, 3440 AA Woerden
Houttuinlaan 4, 3447 GM WOERDEN
Telefoon nr.: 03480 - 77911

I

De conclusie dat het zoogdier Ldh-C gen is ontstaan door duplicatie van het Ldh-A gen (dit proefschrift) is in overeenstemming met de bevinding dat bij de mens deze beide genen op chromosoom 11 gelegen zijn, terwijl het Ldh-B gen op chromosoom 12 wordt gevonden.

- Edwards, Y.H., Povey, S., LeVan, K.M., Driscoll, C.E., Millan, J.L. & Goldberg, E. (1987) Dev. Genet. 8, 219-232.
- Li, S.S.-L., Luedemann, M., Sharief, F.S., Takano, T. & Deaven, L.L. (1988) Cytogenet. Cell Genet. 48, 16-18.

II

In het licht van recente publicaties mag verwacht worden dat ook de 35 kDa eiwitketen, waargenomen in de β-crystalline-fractie uit de ooglens van Camelus dromedarius, een enzym-gerelateerd crystalline is.

- Mulders, J.W.M., Hendriks, W., Blankesteijn, W.M., Bloemendal, H. & de Jong, W.W. (1988) J. Biol. Chem. 263, 15462-15466.
- Borras, T. & Rodokanaki, A. (1988) Proc. Int. Soc. Eye Res. 5, 39.
- Duhaiman, A.S. (1988) Curr. Eye Res. 7, 871-876.

III

Chiou en medewerkers suggereren ten onrechte dat de verwantschap tussen eendehart LDH-B4 en eendelens ε-crystalline, zoals beschreven door Wistow et al., alleen gebaseerd is op partiële sequentievergelijking.

- Chiou, S.-H., Chang, W.-P. & Lin, H.-K. (1988) Biochim. Biophys. Acta 957, 313-317.
- Wistow, G.J., Mulders, J.W.M. & de Jong, W.W. (1987) Nature 326, 622-624.

IV

Het aandikken van het signaal van het βB1-antiserum op de Western blot van kippelens polypeptiden in een recente publicatie van Brahma was niet nodig en komt niet ten goede aan de wetenschappelijke integriteit.

- Brahma, S.K. (1988) Exp. Eye Res. 47, 507-510.

V

Het poneren van een biologische rol voor prostaglandines in de lens als interessante mogelijkheid, enkel gebaseerd op een sequentiegelijkenis tussen prostaglandine F synthase en een lenseiwit dat tot nog toe alleen bij de kikker gevonden is, komt weinig zinnig over.

- Watanabe, K., Fujii, Y., Nakayama, K., Ohkubo, H., Kuramitsu, S., Kagamiyama, H., Nakanishi, S. & Hayaishi, O. (1988) Proc. Natl. Acad. Sci. USA 85, 11-15.

VI

Dat Ortwerth en Olesen het voorkomen van grote hoeveelheden van glutathion-S-transferase-gerelateerde eiwitten in de lenzen van de pijl-inktvis niet bespreken, moet als een omissie worden beschouwd.

- Ortwerth, B.J. & Olesen, P.R. (1988) Exp. Eye Res. 47, 737-750.
- Wistow, G. & Piatigorsky, J. (1987) Science 236, 1554-1556.

VII

Chang en medewerkers wekken ten onrechte de illusie dat zij in een eerdere publicatie de sequentie van het ßs-gen van de karper hebben gepresenteerd.

- Chang, T., Jiang, Y.-J., Chiou, S.-H. & Chang, W.-C. (1988) Biochim. Biophys. Acta 951, 226-229.

VIII

Bij het toepassen van de multivariate middelwaarde-stelling suggereert Rothenberg ten onrechte dat α^* onafhankelijk is van k.

- Rothenberg, T.J. (1971) Econometrica 39, 577-591.

IX

De kwaliteit van voetbal als kijksport kan aanzienlijk verhoogd worden door net als bij basketbal spelregels te introduceren die aan de opbouw en de uitvoering van een aanval een maximale tijdsduur verbinden.

X

De recente uitbreiding van het bier-assortiment van de grote Nederlandse brouwerijen heeft niet geleid tot een verrijking van het aanbod voor de consument.

XI

Het systeem van aftrekposten in het Nederlandse belastingstelsel werkt onrechtvaardig.

XII

Psychofarmaca zijn de spanzeilen van deze tijd.

Nijmegen, 16 februari 1989
Wiljan Hendriks