

Organization of the Human ζ -Crystallin/Quinone Reductase Gene (CRYZ)

PEDRO GONZALEZ, P. VASANTHA RAO, AND J. SAMUEL ZIGLER, JR.¹

Laboratory of Mechanisms of Ocular Diseases, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892

Received September 16, 1993; revised January 28, 1994

ζ -Crystallin is a protein highly expressed in the lens of guinea pigs and camels, where it comprises about 10% of the total soluble protein. It has recently been characterized as a novel quinone oxidoreductase present in a variety of mammalian tissues. We report here the isolation and characterization of the human ζ -crystallin gene (CRYZ) and its processed pseudogene. The functional gene is composed of nine exons and spans about 20 kb. The 5'-flanking region of the gene is rich in G and C (58%) and lacks TATA and CAAT boxes. Previous analysis of the guinea pig gene revealed the presence of two different promoters, one responsible for the high lens-specific expression and the other for expression at the enzymatic level in numerous tissues. Comparative analysis with the guinea pig gene shows that a region of ~2.5 kb that includes the promoter responsible for the high expression in the lens in guinea pig is not present in the human gene. © 1994

Academic Press, Inc.

INTRODUCTION

The ocular lens is a transparent organ characterized by the presence of very high concentrations of proteins called crystallins. These proteins are presumed to be the major determinants of the refractile properties and transparency of the lens. In addition to the α , β , and γ crystallin families, which are present in the lenses of all vertebrates, a number of other crystallins have been found to be present in high amounts in lenses from phylogenetically restricted groups (for reviews, see de Jong *et al.*, 1989; Piatigorsky and Wistow, 1989). Most of these "taxon-specific" crystallins are pyridine nucleotide-dependent oxidoreductases that are also present at enzymatic levels in nonlenticular tissues.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. L31521, L31522, L31523, L31524, L31525, L31526, and L31527.

¹ To whom correspondence should be addressed at Laboratory of Mechanisms of Ocular Diseases, National Eye Institute, National Institutes of Health, Building 6, Room 237, Bethesda, MD 20892. Telephone: (301) 496-6669. Fax: (301) 496-1759.

The acquisition of this new function as a lens crystallin generally occurs without gene duplication and apparently without affecting the catalytic role of the enzyme (Wistow and Piatigorsky, 1987).

ζ -Crystallin/quinone reductase was initially described as a major protein in the lens of the guinea pig (Huang *et al.*, 1987), in which a mutation in the gene is associated with hereditary cataracts (Rodriguez *et al.*, 1992). It was later found to be also present in high amounts in the lens of camels (Garland *et al.*, 1991) and at enzymatic levels in a number of nonlenticular tissues of various species (Rao and Zigler, 1992). As is commonly found with taxon-specific crystallins, only one gene is responsible for the expression at both enzymatic and crystallin levels in guinea pig (Borrás *et al.*, 1990). ζ -Crystallin is distantly related to the alcohol dehydrogenase enzyme family (Rodokanaki *et al.*, 1989); however, a protein from *Escherichia coli* that, despite the great evolutionary distance, has higher sequence similarity with ζ -crystallin has recently been described (N. E. Dixon and P. E. Lilley, Swiss Protein Data Bank P28304). The precise enzymatic role of ζ -crystallin is not known, but analysis of its catalytic properties revealed that while it lacks alcohol dehydrogenase activity, it exhibits an NADPH-dependent quinone reductase activity with characteristics clearly distinct from those of any other known quinone reductase (Rao *et al.*, 1992). This type of activity suggests a possible role as a detoxifying enzyme.

Analysis of the 5' region of the ζ -crystallin gene in guinea pig revealed that the high expression in the lens of this species is mediated by an alternative lens-specific promoter distinct from that responsible for expression in other tissues (Lee *et al.*, 1993; Gonzalez *et al.*, 1994). This is the first case of an enzyme-crystallin in which the acquisition of a new function as a lens protein resulted from the acquisition of a second, alternative promoter.

In this study we present the genomic organization of the human ζ -crystallin/quinone reductase gene as well as a pseudogene lacking introns and compare the 5' region with that of the guinea pig gene. This informa-

TABLE 1
Intron-Exon Junctions of the Human ζ -Crystallin Gene

Exon	Exon length (bp)	5' Splice donor	Intron length	3' Splice acceptor	Codon phase
I	>52	TTGgtaagtggtttttaa Gln	4.7 kb ttttttcttccttagTAT Val	—
II	124	CAGgtagaaaataaaatta Lys	~1.2 kb acattgtcattgcagGTT Lys	0
III	153	AAGgtattgcatttttaa Se	~3.8 kb tatctcttcccttagAAA r	0
IV	164	CAGgtaataacacattct Gly	~4.2 kb tgctttctcttgagTGC Val	II
V	52	GCAGtaagtattttttt Lys	~3.8 kb atatttgttttcagGTT Lys	0
VI	150	AAGgtaaaattttccact Ile	~2.5 kb tcttttattatagAAG Val	0
VII	102	ATAgttagtattcctt Lys	108 bp tgtttttcatggttagGTT Glu	0
VIII	96	AAGgtagaaaagaagta	441 bp ttttcttttcttagGAG	0
IX	>958	TAATGTGTAATAAAGTCT end			

tion should be useful in elucidating the evolutionary mechanisms of recruitment of enzymes as crystallins and their possible roles in the lens and other organs.

MATERIALS AND METHODS

Genomic library screenings. A human genomic library from Promega (Cat. c2091) constructed in λ GEM-11 was screened using nitrocellulose filters (Schleicher & Schuell). Hybridizations were performed at 68°C in 6 \times SSC, 0.5% SDS, 0.5% Denhardt's solution with 100 μ g/ml salmon sperm DNA (Sigma) with 10⁶ dpm/ml of the probe, with washes at 68°C in 2 \times SSC, 0.5% SDS. For the first screening a probe was generated by nested polymerase chain reaction (PCR)² corresponding to nt 455–943 from the human cDNA clone HL103. Further screenings of approximately 0.8 \times 10⁶ pfu were carried out with a probe corresponding to nt 17–121 using the same library and a second one from Clontech (Cat. HL1006d) constructed in EMLB-3 to obtain clones with the 5' region of the gene.

Subcloning and DNA sequencing. Phage DNA was isolated from λ clones, digested with different restriction enzymes, and ligated into previously digested and dephosphorylated pBluescript IKS⁺ (Stratagene). The sequencing reactions were carried out using the Sequenase system (USB) based on the dideoxy chain termination method (Sanger *et al.*, 1977) with 2'-deoxyadenosine 5'-[α -³⁵S]-thio]-triphosphate. The intron/exon junctions and the putative promoter were sequenced in both strands. The complete intron 1 was sequenced in one direction.

The computer comparisons of the sequences were performed using the GCG package in the CONVEX unit at the NIH.

Polymerase chain reaction analysis of the intron sizes. With the exception of introns 1, 7, and 8, which were completely sequenced, the size of the introns was determined by PCR with appropriate oligonucleotides in the exon sequences. To amplify fragments of the sizes expected for the larger introns, we used the buffer described by Ponce and Micol (1992). This buffer has been specifically developed to

get amplifications up to 6 kb. The *Taq* polymerase and all other reagents used were purchased from Perkin-Elmer.

Rapid amplification of cDNA 5' end (RACE). Total RNA was isolated from 1 g of normal human liver using an RNA extraction system (Stratagene) based on the procedure of Chomczynski and Sacchi (1987). The first strand cDNA synthesis was performed by priming the total RNA with the oligo (5'-AAGCCATAAGCTCTAGCAAT-3') antisense from ζ -crystallin cDNA using MLV reverse transcriptase (BRL). DNA was tailed with TdT from BRL. The first PCR amplification was performed using a poly(T) oligonucleotide with an *Eco*RI restriction site (5'-GGACTCGAATTCGACTGCTTTT-TTTTTT-TTTTT-3') and the internal ζ -crystallin antisense oligo (5'-CCA-GAGATCGTCTGCTAGT-3'). After 10 cycles with an annealing temperature of 37°C, 0.2 μ l of the reaction mixture was used for a second amplification with the oligonucleotide 5'-GGACTCGAATTCGACTGCT-3' and a ζ -crystallin antisense oligo slightly modified in the 5' to generate an *Eco*RI site (5'-GTGGTTGAATTCTGTAAGTAC-CAGAGCGAATGTATGTCT-3'). Amplification products were digested with *Eco*RI and subcloned in pBluescript IKS⁺ (Stratagene) for sequencing.

Primer extension analysis. A 36-mer synthetic oligonucleotide (5'-CTAGAGTGGGAATTAATAATCTGCGTGGGCTTCT-3') complementary to the first untranslated exon of ζ -crystallin was end-labeled with [³²P]ATP using T4 polynucleotide kinase (Stratagene). The unincorporated nucleotides were removed from the ³²P-labeled oligonucleotide using a NICK-column (Pharmacia), and 5 \times 10⁶ cpm of this probe was hybridized with 50 μ g of total RNA from human liver with reverse transcriptase buffer, 4 U of RNasin (Promega), and 200 U/ μ l AMV Reverse transcriptase (BRL). After 2 h of incubation at 42°C, the extension products were phenol extracted and ethanol precipitated prior to electrophoretic separation in a 6% acrylamide/7 M urea gel for analysis.

Southern blotting. Genomic DNA from human liver was extracted as described in Sambrook *et al.* (1989). DNA was transferred to nylon membranes (NYTRAN, Schleicher & Schuell). Hybridizations were performed with the same stringency as that for the library screening. Two fragments generated by nested PCR from the human liver cDNA clone HL103 insert corresponding to nt 17–121 and 1133–1427 were used to identify the bands corresponding to the 5' and 3' ends of the gene, respectively. The insert of clone HL103 was used as a full-length probe. The analysis of the clones was performed following basically the same method as that used for genomic DNA.

² Abbreviations used: ADH, alcohol dehydrogenase; ARE, antioxidant responsive element; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; SDS, sodium dodecyl sulfate; SSC, 0.015 M NaCl/0.015 M Na₃ citrate, pH 7.0; HSE, heat shock element; nt, nucleotide.

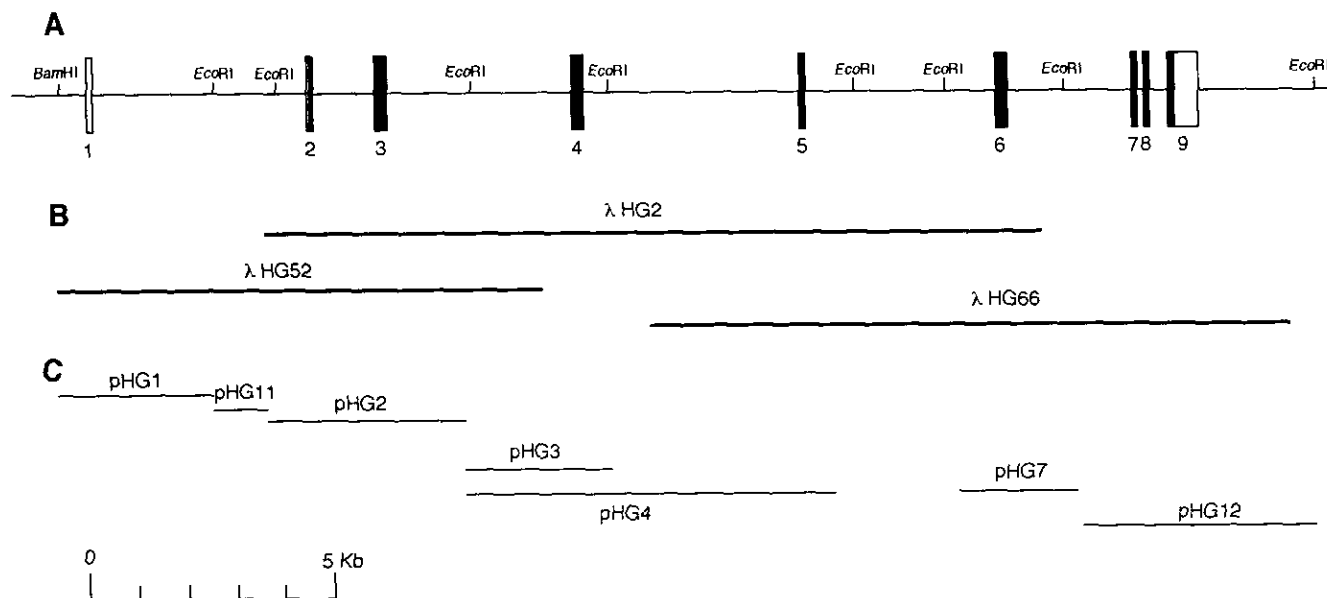


FIG. 1. (A) The structure of the human ζ -crystallin gene. The exons are shown as boxes and numbered below. Solid boxes designate coding regions, and open boxes represent the 5' and 3' untranslated regions. The locations of the restriction sites used to subclone different fragments of the λ -clones are also indicated. (B) The relative positions of the λ -genomic clones. (C) The fragments subcloned into pBluescript KSII⁺ that were used for the sequence analysis.

RESULTS

Intron-Exon Structure of the Human ζ -Crystallin Gene

The extent of the genomic clones used in this study and the structure of the human ζ -crystallin gene are shown in Fig. 1. The gene spans about 20 kb and includes one untranslated and eight translated exons. No differences were found between the exon sequences in the genomic clones and the previously reported cDNA sequence (Gonzalez *et al.*, 1993). The ATG start codon is located 14 bp into exon 2. The last exon, exon 9, includes 53 amino acids of the C-terminal region of the protein plus the 3' noncoding region. The cDNA clone previously reported ended only four nucleotides after a consensus polyadenylation signal (AATAAA), and no poly(A) tail was found; therefore, the exact size of exon 9 is not completely defined since other consensus polyadenylation sequences are also present in the 3' region of the gene. However, the analysis of cDNA clones from guinea pig (Rodokanaki *et al.*, 1989) showed that a functional polyadenylation signal is located in the same relative position as that found in the human cDNA clone. All the intron/exon boundaries have the consensus sequences GT and AG (Breathnach and Chambon, 1981; Mount, 1982). The lengths of introns 2, 3, 4, 5, and 6 were estimated based on PCR amplifications between exons and the restriction pattern of the genomic clones. Introns 7 and 8 are the smallest, and their sizes (108 and 441 nt, respectively) were determined by direct sequencing. The intron/exon structure and splice sequences of the human ζ -crystallin gene are summarized in Table 1.

Amplification of the cDNA 5' End in Human ζ -Crystallin and Primer Extension Analysis

A RACE experiment was performed with total RNA from human liver to complete the sequence of the 5' end of the ζ -crystallin cDNA. Four clones (RaceHL1,2,3,4) were tested by internal PCR with specific oligonucleotides from the 5' region of ζ -crystallin; two of them (RaceHL3,4) showed positive amplification and were sequenced. Clones RaceHL3 and RaceHL4 extended 34 and 55 nucleotides, respectively, upstream of the cDNA sequence previously described (Gonzalez *et al.*, 1993). Further analysis of the genomic clones showed that this sequence corresponds to an untranslated exon present

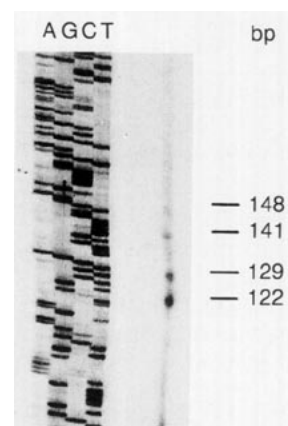


FIG. 2. Primer extension mapping of the 5' end of human ζ -crystallin in liver. Lanes A, G, C, and T represent an unrelated sequencing reaction used as a size marker. The lane on the right shows the products of the primer extension obtained as described under Materials and Methods. The sizes of the extended products corresponding to the more intense bands are also shown.

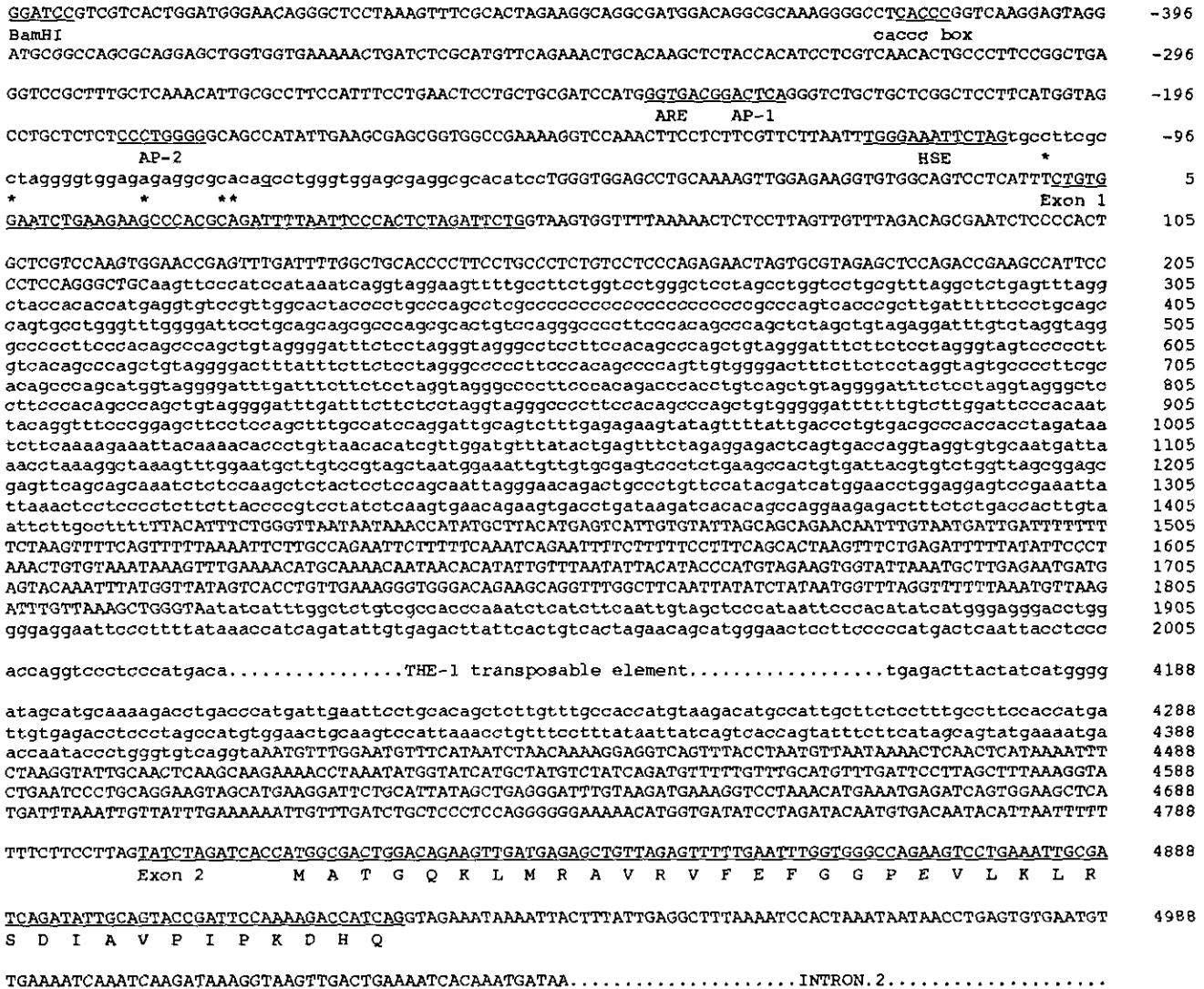


FIG. 3. Sequence corresponding to the 5' region of the human ζ-crystallin. For convenience, nucleotides are numbered starting with the first nucleotide of the longest RACE product from liver RNA. The locations of the putative transcription start points according to the primer extension experiment are indicated with asterisks below the corresponding nucleotide. The exon sequences are underlined, and the amino acids are shown below the corresponding codons. The fragments not showing similarity with the corresponding guinea pig sequences (Gonzalez *et al.*, 1994) are indicated with lowercase letters. Sequence from nucleotides 1823 to 4411 corresponds to the THE-1 transposable element.

approximately 4.5 kb upstream of the first translated exon (Exon 2).

A primer extension experiment with human liver RNA (Fig. 2) shows the presence of several putative transcription start points located 75 to 101 nucleotides upstream of the first nucleotide of the longest product of the RACE experiment. Multiple transcription start points are frequently observed in housekeeping promoters.

The 5' Region of the Human ζ-Crystallin

The sequence located upstream of the first exon is rich in G+C and lacks TATA and CAAT boxes. Some sequences similar to putative regulatory elements,

CACCC box, antioxidant responsive element (ARE) (Rushmore *et al.*, 1991), AP-1, AP-2, and heat shock element (HSE) (Amin *et al.*, 1988), are found in the human gene (Fig. 3). The overall similarity between the two species in this region (>80%) is significantly higher than the average 60% similarity found in non-functional intron sequences, a finding consistent with its putative role as promoter.

The high expression of ζ-crystallin in the lens of guinea pig is mediated by a virtually lens-specific promoter located between the liver untranslated exon and exon 2 (Lee *et al.*, 1992). To obtain some insight into the possible mechanism of recruitment of ζ-crystallin as a lens protein in guinea pig, the human intron 1 was completely sequenced and compared with the cor-

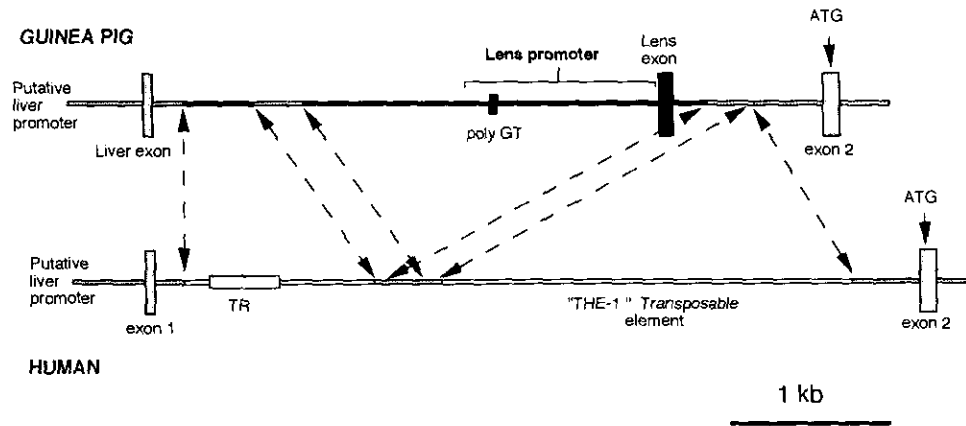


FIG. 4. Comparison of the 5' region of the human ζ -crystallin gene with the guinea pig (Gonzalez *et al.*, in preparation). The regions conserved (similarity > 60%) between the two species are presented as gray boxes. The sequences specific to the human or guinea pig genes are presented as white and black boxes, respectively. The human THE-1 repetitive element is also shown. The region labeled as lens promoter in the guinea pig gene is based on the functional analysis by Lee *et al.* (1994).

responding region in the guinea pig gene (Gonzalez *et al.*, 1994).

There were three main areas in which the sequence in one species did not show homology with the other (Fig. 4). The first of these areas starts 170 bp downstream of the first exon and affects a fragment of 1.2 kb in human and 470 bp in guinea pig that cannot be aligned. A tandemly repeated sequence of about 50 bases with 10 repeats is located in this area in the human gene.

A second difference corresponds to a sequence in the human gene identified as a member of the transposon-like family described by Paulson *et al.* (1985) that spans 2.6 kb within the first intron, ending 400 bp upstream of exon 2. This sequence contains two flanking long terminal repeats and shows similarity with the THE-1 family of transposable elements (Paulson *et al.*, 1985; Finnegan, 1989). This family of transposon-like elements has been reported in the human genome (where approximately 10,000 copies are estimated to be present) and in the monkey but not in the rodent genome and must therefore be considered a relatively recent acquisition in the evolution of the ζ -crystallin gene. These elements are able to transpose by their *de novo* insertion into other genes (Deka *et al.*, 1988). The relative position of the insertion and the small effect on the original sequence (only 34 nt are missing with respect to the guinea pig intron sequence) would suggest that it probably has not affected the expression of the gene.

The more interesting difference between the human and the guinea pig genes involves a fragment of ~2.5 kb of the guinea pig gene that is not present in the human intron. This guinea pig fragment includes the lens-specific exon plus 220 nucleotides downstream and 2.2 kb upstream of it. This includes the complete region characterized as the lens-preferred promoter responsible for the high expression of ζ -crystallin in guinea pig (Lee *et al.*, 1992). A fragment of 300 nt situ-

ated upstream of the THE-1 transposable element in the human gene shares similarity with sequences located in both flanking areas of the 2.5-kb fragment containing the guinea pig lens promoter. However, these two guinea pig sequences show only a very weak similarity with each other.

Structure of the ζ -Crystallin Pseudogene

Figure 5 shows the sequence of the human intronless gene. The open reading frame has been interrupted at different points, indicating that this is not an active gene, but a processed pseudogene (Vanin, 1984). The similarity with the human cDNA (Gonzalez *et al.*, 1993) is 81%. Since the rate of nucleotide substitutions for pseudogenes has been calculated by Li *et al.* (1981) as 4.6×10^{-9} substitutions per nucleotide site per year, the formation of the pseudogene can be estimated to have occurred about 40 million years ago. According to this estimation this pseudogene would not be expected to be present in rodents, a prediction consistent with the observed absence of ζ -crystallin pseudogenes in guinea pig (Borrás *et al.*, 1990). The existence of this pseudogene indicates that the gene is expressed in the germinal cell line consistent with expression at the level of that of a housekeeping enzyme. It is also interesting to note that the similarity in the 5' region of the pseudogene with the products of RACE from human mRNA and with the genomic sequences helps to confirm the 5' untranslated sequence obtained in the RACE experiment.

Southern Blot Analysis

Figure 6 shows the results of two Southern blots of human liver DNA digested with *EcoRI*, *HindIII*, and *PstI* and hybridized separately with two short probes corresponding to sequences located in exons 2 and 9 of the human gene, respectively. Only one band hybridized with high intensity in each lane; however, one

```

accttggcctggcctccagcacccttagattaactgaaccccaactccaaaagtcttgttgctgcaaagttagaagcagc 81
                                     M A
ctggccagcctccttagtggaatagtaaggagcccttcagttttgatcccactctagagcttttatctagtgaccATGGCA 162
T E Q K L V R A V R V F E F G R P E V L K L S Q M L Q
ACTGAACAAAAGCTGGTGAGAGCTGTTAGAGTGTGAAATTGGTAGACCAGAAGTCTGAAACTCAGTCAGATGTTGCAG 243
Y Q F Q K T I S V I R F M R V V S T P W R H T F T L V
TACCAGTTCAAAAGACCATTAGTGTAATAAGGTTTCATGCGTGTGGTGTCAACCCCATGGAGACATACATTTACTCTGGTA 324
F N R K P P Y S I L R C N W G N G S C W R E C I Y F Q
TTAATAGAAAAACACCTTATTCTATTCTCAGATGTAACGGGGTAACGGAAGCTGTGGAGGGAATGCATCTACTTTCAA 405
E K W T V F T T G T I S G S R L R C F T N C L K N W T
GAAAAGTGGACAGTTTTCTACTACCGGCACAATCTCTGGGAGCAGATTACGCTGTTTTACCAACTGCCTGAAAAACTGGACT 486
L N K M L P L V S H I L L H I E L * F T V S G M K A G
TTAAACAAAATGCTGCCATTGGTATCCCATATCTTACTGCATATCGAGCTCTGATTACAGTGTCTGGTATGAAAGCTGGA 567
E F W F M G L V E E L Q * Q H A K L L E L C L E G F G
GAGTTCGGTTCATGGGGCTAGTGGAGGAGCTGCAGTAGCAACATGCCAAATGCTAGAGTTATGTCTTGAAGTTTTGGG 648
H S W Y * G T K R L L Q N G A H K V F N Y K E V N Y V
CACAGTGGTACTGAGGGACAAAAGACTTTTGCAAAATGGAGCCACAAAAGTGTTAATTACAAAGAAGTAACTATGTT 729
D K I K K C V I R R I D V I L K * * L N K S * * I C E
GATAAAATTAAGAAATGTGTGATAAGAAGGATTGATGTGATATGAAATAATAGCTAAATAAATCTTAGTAAATATGCGAA 810
L L Y M E D K * * V L A T V V P L K * T H G T L W Q K
TTGTTGTATATGGAAGATAAGTAATAGGTGTTGGCAACAGTGGTACCATTGAAATAAACCCATGGGACACTATGGCAAAAAG 891
S L V K L S L L Y F P Q P R R N L S N L Q Q P V K L E
AGTCTAGTAAATTTGCTTACTCTATTTTCTCAACCAAGGAGGAATCTCAGCAATTTGCAGCAGCCTGTCAAGCTGGAA 972
W K L V G * N L * * V L T I C W R G G P G S * K Y L S
TGGAAATGGTGGCTGAAACCTGTAATAAGTTCTCACTATATGTTGGAGAGGTGGCCAGGCTCATGAAAATATCTCTCA 1053
R Q W G * C K N D S S L L
CGGCAATGGGGCTAGTGTAATAATGATTCTTCTTTTGATgattaattctttcatggatttcagaggttgtcttttgccc 1134

agttttagttacacaatctttgtttaattaacattcatttgatccagtgagtttcttatgggaaaaaagtaagatttttc 1215

```

FIG. 5. Structure of human ζ -crystallin pseudogene. The sequence similarity between the functional gene and the pseudogene starts at the position indicated by an arrow. The open reading frame is interrupted by numerous stop codons labeled with asterisks. The presence of these stop codons indicates the lack of functionality of the pseudogene.

additional weaker band can be observed, especially in the blot hybridized with the 5' probe. The sizes of the stronger bands are the same as those of the corresponding restriction fragments of the genomic clones from the functional gene; the weaker bands match those of the intronless pseudogene.

DISCUSSION

The main purpose of this study was to isolate and characterize the human ζ -crystallin gene to better understand its function and regulation of expression. In addition, the data should provide some insight into the mechanism of recruitment of this enzyme as a lens crystallin through the comparative analysis of the human gene with those of other species.

The RACE experiment revealed a new untranslated exon not found in the original analysis of the human cDNA (Gonzalez *et al.*, 1993). The sequence upstream of this exon has features of a housekeeping promoter

(G+C rich, lack of TATA and CAAT box). Except for a CACCC box and a sequence similar to the HSE, there are no clear consensus regulatory sequences conserved with respect to the guinea pig housekeeping promoter (Gonzalez *et al.*, 1994). However, the higher degree of conservation of this region suggests that it is functional as a promoter. The presence of this type of promoter is also consistent with a function for ζ -crystallin as a housekeeping enzyme, perhaps involved in a detoxification process as previously proposed (Rao *et al.*, 1992).

One of the more interesting features of the ζ -crystallin gene is the great difference in its expression in lenses from different species. This is especially relevant in the comparative analysis of the region corresponding to the first intron, where a lens-specific promoter is present in the guinea pig gene (Lee *et al.*, 1993; Gonzalez *et al.*, 1994). Comparison of the sequences of the first intron from the human and the guinea pig genes showed three main differences. Two of them (the translocation in the 5' region of the intron and the insertion

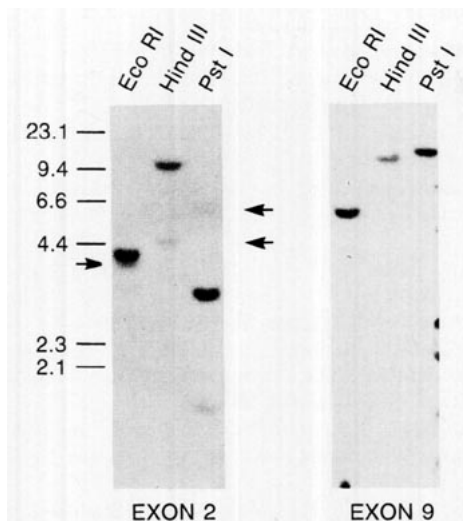


FIG. 6. Southern blot analysis of the human ζ -crystallin gene. Human genomic DNA digested with the indicated enzymes was probed separately with two PCR fragments (1×10^6 cpm/ml) corresponding to exon 2 (left) and exon 9 (right). The pattern of the more intensely hybridizing bands is consistent with that of the genomic clones of the functional gene. The pattern of weaker bands (indicated with arrows) matches that of the intronless pseudogene. The lack of conservation between the gene and pseudogene sequences in the untranscribed region of exon 9 explains the absence of weaker bands in the blot hybridized with the 3' probe.

of the THE-1 family transposable element in the human gene) apparently are not implicated in causing the difference of expression in the lens; they affect areas for which there is no evidence of a functional role, especially with regard to the lens expression in guinea pig. The difference between the human and the guinea pig genes that accounts for the different level of expression in the lens is the 2.5-kb fragment of guinea pig that includes the lens promoter and is missing in the human gene. Analysis of the guinea pig lens promoter has shown that this fragment contains all the regulatory elements necessary for high lens expression, and there is no evidence for regulatory elements outside of this area (Lee *et al.*, 1994). This result raises the possibility that the recruitment of ζ -crystallin/quinone reductase as a lens crystallin occurred by the insertion of this 2.5-kb fragment into the first intron of the original gene. However, it could also be interpreted as a deletion in the evolution of the human gene. In fact, the observation that the two guinea pig sequences flanking the fragment that is missing in the human gene share some similarity with each other and with only one fragment of the human sequence in that region is probably more consistent with a deletion of the guinea pig fragment rather than an insertion of a new one.

The analysis of the human ζ -crystallin thus does not provide a definitive answer to the question of whether the lens promoter in the guinea pig gene (Lee *et al.*, 1993) is an ancestral feature of the gene that was eliminated in most evolutionary lineages or whether it is a new acquisition in guinea pig. However, the observa-

tion that the complete region corresponding to the guinea pig lens promoter is absent in the human gene suggests that it was not originally involved in the regulation of the expression of the gene at enzymatic levels since this constitutive expression of ζ -crystallin is apparently unaffected. Therefore, while a role for neutral evolution cannot be ruled out, these results suggest that the recruitment of ζ -crystallin as a lens protein in guinea pig did not occur through modification of regulatory elements present in the original gene, as has been proposed as a general mechanism for the neutral recruitment of other enzyme/crystallins (De Jong *et al.*, 1989). The recruitment of ζ -crystallin, and perhaps other pyridine nucleotide-dependent oxido-reductases as taxon-specific crystallins, may reflect some beneficial effect of the overexpression of these proteins in the lens that provides selective advantage. In this sense it is interesting to note that due to the ability of ζ -crystallin to bind NADPH, the high concentration of this protein in the lenses of guinea pigs and camels results in an extremely high level of NADPH as well (Rao and Zigler, 1990). A similar increase in the concentration of reduced pyridine nucleotides has been found in the lenses of species with other nucleotide-binding enzyme-crystallins (Zigler and Rao, 1992). The fact that synthesis and maintenance of these very high nucleotide levels require the expenditure of energy may suggest that either the enzyme/crystallins themselves or the high concentration of nucleotides associated with them provides some definite advantage to the lens. It has been proposed (Rao and Zigler, 1992) that increased levels of NADH or NADPH may play a protective role against oxidation in the lens, which is believed to be a major factor in the development of cataracts (Spector, 1991). It is especially interesting that one of the species that has high levels of ζ -crystallin in the lens, the camel (*Camellus dromedarius*), lives in a particularly hostile environment, where it is exposed to extensive UV radiation. Apparently, despite such exposure, cataracts are not commonly found even at older ages (34–37 years) (K. F. Tabbara, Saudi Arabia, pers. comm., 1993). A study of the ζ -crystallin gene in another camelid (llama, *Llama guanacoe*) that also shows high expression of this protein in the lens is currently in progress in our laboratory. Comparative analysis of the regulatory regions of the llama gene with those presently available should provide new insights into the evolution of the gene and its role in the lens and other tissues.

REFERENCES

- Amin, J., Ananthan, J., and Voellmy, R. (1988). Key features of heat shock regulatory elements. *Mol. Cell. Biol.* **8**: 3761–3769.
- Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**: 831–839.
- Borrás, T., Jörnvall, H., Rodokanaki, A., Gonzalez, P., Rodriguez, I., and Hernández-Calzadilla, C. (1990). The transcripts of zeta-Crystallin, a lens protein related to the alcohol dehydrogenase

- family, are altered in a guinea pig hereditary cataract. *Exp. Eye Res.* **50**: 729–735.
- Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162**: 831–839.
- De Jong, W. W., Hendriks, W., Mulders, J. W. M., and Bloemendal, H. (1989). Evolution of the eye lens crystallins: The stress connection. *Trends Biochem. Sci.* **14**: 365–368.
- Deka, N., Wong, E., Matera, A. G., Kraft, R., Leinwand, L. A., and Schmid, C. W. (1988). Repetitive nucleotide sequence insertions into a novel calmodulin related gene and its processed pseudogene. *Gene* **71**: 123–134.
- Finnegan, D. J. (1989). Eukaryotic transposable elements and genome evolution. *Trends Genet.* **5**: 103–107.
- Garland, D., Rao, P. V., Del Corso, A., Mura, U., and Zigler, J. S., Jr. (1991). ζ -Crystallin is a major protein in the lens of *Camellus dromedarius*. *Arch. Biochem. Biophys.* **285**: 134–136.
- Gonzalez, P., Hernández-Calzadilla, C., Rao, P. V., Rodriguez, I. R., Zigler Jr., J. S., and Borrás, T. (1994). Comparative analysis of the ζ -Crystallin/Quinone reductase gene in guinea pig and mouse. *Mol. Biol. Evol.* **11**: 305–315.
- Gonzalez, P., Rao, P. V., and Zigler, J. S., Jr. (1993). Molecular cloning and sequencing of the zeta-crystallin/quinone reductase cDNA from human liver. *Biochem. Biophys. Res. Commun.* **191**: 902–907.
- Huang, Q. L., Russell, P., Stone, S. H., and Zigler, J. S., Jr. (1987). ζ -Crystallin, a novel lens protein from the guinea pig. *Curr. Eye Res.* **6**: 725–732.
- Lee, D. C., Gonzalez, P., Rao, P. V., Zigler, J. S., Jr., and Wistow, G. J. (1993). Carbonyl-metabolizing enzymes and their relatives recruited as structural proteins in the lens. In "Enzymology and Molecular Biology of Carbonyl Metabolism" (H. Weiner, Ed.), pp. 159–168, Plenum, New York.
- Lee, D. C., Gonzalez, P., and Wistow, G. (1994). ζ -Crystallin: A Lens-specific promoter and the gene recruitment of an enzyme as a crystallin. *J. Mol. Biol.* **236**: 669–678.
- Li, W. H., Gojobori, T., and Nei, M. (1981). Pseudogenes as a paradigm of neutral evolution. *Nature* **292**: 237–239.
- Mitchell, P. J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**: 371–378.
- Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**: 459–472.
- Novacek, M. J. (1992). Mammalian phylogeny: Shaking the tree. *Nature* **356**: 121–125.
- Paulson, K. E., Deka, N., Schmid, C. W., Misra, R., Schindler, C. W., Rush, M. G., Kadyk, L., and Leinwand, L. (1985). A transposon-like element in human DNA. *Nature* **316**: 359–361.
- Piatigorsky, J., and Wistow, G. J. (1989). Enzyme/crystallins: Gene sharing as an evolutionary strategy. *Cell* **57**: 197–199.
- Ponce, M. R., and Micol, J. L. (1992). PCR amplification of long DNA fragments. *Nucleic Acids Res.* **20**: 623.
- Rao, Ch. M., and Zigler, J. S., Jr. (1992). Levels of reduced pyridine nucleotides and lens photodamage. *Photochem. Photobiol.* **56**: 523–528.
- Rao, P. V., and Zigler, J. S., Jr. (1990). Extremely high levels of NADPH in guinea pig lens: Correlation with ζ -crystallin concentration. *Biochem. Biophys. Res. Commun.* **167**: 1221–1228.
- Rao, P. V., Krishna, C. M., and Zigler, J. S., Jr. (1992). Identification and characterization of the enzymatic activity of ζ -crystallin from guinea pig lens. *J. Biol. Chem.* **267**: 96–101.
- Rao, P. V., and Zigler, J. S., Jr. (1992). Purification and characterization of ζ -crystallin/quinone reductase from guinea pig liver. *Biochem. Biophys. Acta* **1117**: 315–320.
- Rodokanaki, A., Holmes, R. K., and Borrás, T. (1989). Zeta-crystallin, a novel protein from the guinea pig lens, is related to alcohol dehydrogenases. *Gene* **78**: 215–224.
- Rodriguez, I. R., Gonzalez, P., Zigler, J. S., Jr., and Borrás, T. (1992). A guinea-pig hereditary cataract contains a splice site deletion in a crystallin gene. *Biochim. Biophys. Acta* **1180**: 44–52.
- Rushmore, T. H., Morton, M. R., and Pickett, C. B. (1991). The antioxidant responsive element: Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. *J. Biol. Chem.* **266**: 11632–11639.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nickler, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463–5464.
- Spector, A. (1991). The lens and oxidative stress. In "Oxidative Stress: Oxidants and Antioxidants" (H. Sies, Ed.), pp. 529–558, Academic Press, New York.
- Vanin, E. F. (1984). Processed pseudogenes: Characteristics and evolution. *Biochim. Biophys. Acta* **782**: 231–241.
- Wistow, G., and Piatigorsky, J. (1987). Recruitment of enzymes as lens structural proteins. *Science* **236**: 1554–1556.
- Wistow, G., Mulders, W. M., and de Jong, W. W. (1987). The enzyme lactate dehydrogenase as a structural protein in avian and crocodilian lenses. *Nature* **326**: 622–624.
- Zigler, J. S., Jr., and Rao, P. V. (1991). Enzyme/crystallins and extremely high pyridine nucleotide levels in the eye lens. *FASEB J.* **5**: 223–225.