Gene, 133 (1993) 295–299

GENE 07371

Sequence of the human lens βB2-crystallin-encoding cDNA

(Recombinant DNA; structural protein; eye; eye cataract; polymerase chain reaction; protein similarity)

Carolyn Chambers and Paul Russell

Section on Cataracts, Laboratory of Mechanisms of Ocular Diseases, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received by J.A. Engler: 26 January 1993; Revised/Accepted: 26 April/26 April 1993; Received at publishers: 28 June 1993

SUMMARY

This study reports the nucleotide (nt) sequence of the human βB2-crystallin (hβB2-Cry)-encoding cDNA (h*βB2-cry*). The h*βB2-cry* gene encodes a major structural protein in the lens of the vertebrate eye. Sequence information obtained from mouse genomic clones and the mouse *βB2-cry* cDNA (m*βB2-cry*) sequence was used to design oligodeoxyribonucleotide primers, corresponding to exon 1, transcription start points, and termination and polyadenylation sites, that were used in the polymerase chain reaction (PCR) to generate full-length cDNA clones from total human lens RNA. In addition, cDNA libraries were made with λgt11 from both human fetal and adult lens tissue RNAs, and screened with a m*βB2-cry* cDNA clone. The full-length h*βB2-cry* cDNA is 721 bp and contains an open reading frame of 612 nt. It codes for a 23-kDa protein of 205 amino acid residues. Comparison of the overall nt and deduced aa sequences shows a greater similarity of h*βB2-cry* to bovine *βB2-cry* than to m*βB2-cry* or rat *βB2-cry*.

INTRODUCTION

The most abundant water-soluble proteins found in the lens of the vertebrate eye are called crystallins (Cry), and transparency of the lens is thought to be due to the interactions of these various Cry (Benedek, 1971; Bettelheim and Siew, 1982). This optical property is thought to be preserved by the arrangement of the fiber cells with exact spatial distribution of the Cry (Delaye and Tardieu, 1983). A disruption of protein-protein interactions within the fiber cells in the lens could lead to the formation of a cataract. If a Cry has a change in its molecular structure, it may not be able to form aggregates with the other lens Cry. There are four antigenically distinct families of Cry found in the lens of the vertebrate eye (Piatigorsky, 1984). The β-Cry family is the most prevalent comprises about 50% of the water soluble protein. Several members of this family have been cloned and the cDNAs characterized: *βB1-cry* in chick (Hejtmancik et al., 1986), bovine (Quax-Jeuken et al., 1984), and rat lenses (Den Dunnen et al., 1985); *βA3/A1-cry* in bovine (Quax-Jeuken et al., 1984), chick (Peterson and Piatigorsky, 1986), rat (Aarts et al., 1989b), and mouse lenses (Inana et al., 1982); and cDNAs for *βA2-cry* and *βA4-cry* in bovine lens (Van Rens et al., 1991). The cDNA for *βB2-cry* has been studied from rat (Aarts et al., 1989a), mouse (Chambers and Russell, 1991), and bovine lenses (Hogg et al., 1987). These sequences have been studied because of the possibility that the alteration of these sequences in some animals are responsible for lens opacities.

There are several animal models for inherited cataracts that show a possible correlation between a defective Cry and cataract formation. One such model is the Philly

Correspondence to: Dr. C. Chambers, LMOD, Bldg. 6, Rm. 228, National Eye Institute, NIH, Bethesda, MD 20892, USA. Tel. (1-301) 496-7471; Fax (1-301) 496-1759; e-mail: CHAMBERS@HELIX.NIH.GOV

Abbreviations: aa, amino acid(s); bp, base pair(s); cDNA, DNA complementary to RNA; b*βB2-cry*, bovine *βB2-cry*; Cry, crystallin; *cry*, gene (DNA) encoding Cry; GCG, Genetics Computer Group (Madison, WI, USA); h*βB2-cry*, human *βB2-cry*; kb, kilobase(s) or 1000 bp; m*βB2-cry*, mouse *βB2-cry*; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; PCR, polymerase chain reaction; r*βB2-cry*, rat *βB2-cry*; *tsp*, transcription start point(s); *UTR*, untranslated region(s).

mouse that has a 12-nt deletion in the β*B2-cry* gene. This results in a deletion of 4 aa in the translated protein (Chambers and Russell, 1991). In the 13/N strain of guinea pig, there is a dinucleotide deletion of the consensus AG acceptor splice site of intron 6 of the taxon specific ζ-Cry. As a result, exon 7 is deleted causing the ζ-Cry polypeptide to be 34 aa shorter in these animals (Rodriguez et al., 1992). In each of these animal models an aberrant or mutant protein is made which may be correlated with cataract formation. There is a Coppock-like cataract in humans that has been linked to the human γ-*cry* gene cluster on chromosome 2 (Lubsen et al., 1987). In that study the exact molecular basis for the cataract was unclear; however, there may be aberrant expression of a γ-*cry* gene (Lubsen et al., 1987).

It has been shown that there are two β*B2-cry* genes in man (Aarts et al., 1987) that are found on chromosome 22. One has been shown to be a pseudogene (Brakenhoff et al., 1992) and the other is expressed. This paper reports the sequencing of the cDNA that codes for the human βB2-Cry protein.

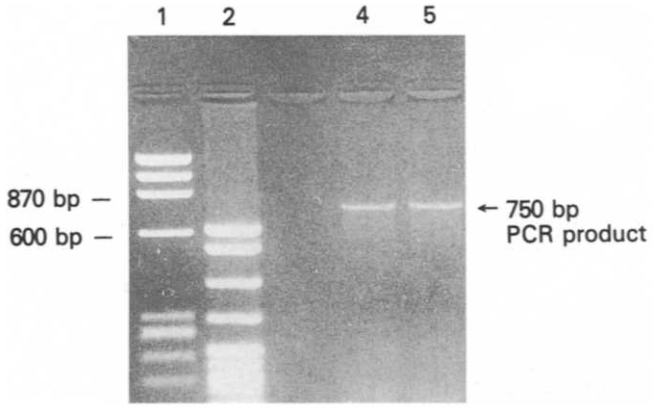

Fig. 1. Agarose gel of PCR-generated product. Lanes: 1 and 2 are DNA markers (φX174 RF DNA/*Hae*III and pBR322 DNA/*Msp*I, respectively); 4 and 5 are PCR products from two different human lens RNA samples with the same reaction conditions. The gel was stained with ethidium bromide. **Methods:** Using a PCR-RNA kit (Perkin-Elmer Cetus, Norwalk, CT, USA) total human lens RNA was reverse transcribed (Moloney Murine Leukemia virus reverse transcriptase) with subsequent PCR amplification (GeneAmp PCR with AmpliTaq DNA polymerase) in the same tube. The following PCR cycling conditions were used: melting at 94°C for 1 min; annealing at 55°C for 1 min; polymerize at 72°C for 2 min; and repeated for 35 cycles. The 5′ oligo (5′-AATTCGCGGCCGCCGGCACTCGCGGGGCTGGCGATCTG) corresponded to a sequence in the first exon (from unpublished mouse genomic sequences) and contained an *Eco*RI-*Not*I adapter at the 5′ end. The 3′ oligo (5′-AATTCGCGGCCGCCACGAGCCACACTTTATTCTTCACTT) corresponded to sequences in the 3′*UTR* (from mouse β*B2-cry* cDNA; Chambers and Russell, 1991) and contained an *Eco*RI-*Not*I adapter at the 5′ end. The 750-bp fragment was excised from the agarose gel, the DNA extracted using Millipore ultrafree-MC filter units, and cloned into the *Eco*RI site of pGEM4Z vector (Promega, Madison, WI, USA). Positive clones were obtained after screening with digoxigenin-UTP-labelled mβ*B2-cry* cDNA clone KSNM1 (Chambers and Russell, 1991) (Genius non-radioactive labeling kit, Boehringer Mannheim, Indianapolis, IN, USA). Several positive clones were sequenced (Sequenase version 2.0, US Biochemicals, Cleveland, OH, USA).

```
GGACAGTCCA CCATGGCCTC AGATCACCAG ACCCAGGCGG GCAAGCCACA GTCCCTCAAC    60
            M  A   S  D  H  Q   T  Q  A   G  K  P   Q  S  L  N     16

CCCAAGATCA TCATCTTTGA GCAGGAAAAC TTTCAAGGCC ACTCGCATGA GCTCAATGGG   120
 P  K  I   I  I  F   E  Q  E  N   F  Q  G   H  S  H   E  L  N  G    36

                                          <-----#1482
CCCTGCCCCA ACCTGAAGGA AACTGGCGTG GAGAAGGCAG GTTCTGTCCT AGTGCAGGCT   180
 P  C  P   N  L  K   E  T  G  V   E  K  A   G  S  V   L  V  Q  A    56

GGACCCTGGG TGGGCTATGA ACAGGCCAAC TGCAAGGCCG AGCAGTTTGT GTTTGAGAAG   240
 G  P  W   V  G  Y   E  Q  A  N   C  K  G   E  Q  F   V  F  E  K    76

GGTGAGTACC CCCGCTGGGA CTCATGGACC AGCAGCCGAA GGACGGACTC CCTCAGCTCC   300
 G  E  Y   P  R  W   D  S  W  T   S  S  R   R  T  D   S  L  S  S    96

CTGAGGCCCA TCAAAGTGGA CAGCCAAGAG CACAAGATCA TCCTCTATGA AAACCCCAAC   360
 L  R  P   I  K  V   D  S  Q  E   H  K  I   I  L  Y   E  N  P  N   116

TTCACCGGGA AGAAGATGGA AATCATAGAT GACGATGTAC CCAGCTTCCA CGCCCATGGC   420
 F  T  G   K  K  M   E  I  I  D   D  D  V   P  S  F   H  A  H  G   136

TACCAGGAGA AGGTGTCATC TGTGCGGGTG CAGAGTGGCA CGTGGGTTCG CTACCAGTAC   480
 Y  Q  E   K  V  S   S  V  R  V   Q  S  G   T  W  V   G  Y  Q  Y   156

CCCGGCTACC GTGGACTGCA GTACCTGCTG GAGAAGGGAG ACTACAAGGA CAGCAGCGAC   540
 P  G  Y   R  G  L   Q  Y  L  L   E  K  G   D  Y  K   D  S  S  D   176

                                                    #1483----->
TTTGGGGCCC CTCACCCCCA GGTGCAGTCC GTGCGCCGTA TCCGCGACAT GCAGTGGCAC   600
 F  G  A   P  H  P   Q  V  Q  S   V  R  R   I  R  D   M  Q  W  H   196

CAACGTGGTG CCTTCCACCC CTCCAACTAG TGCCCTCCCC ACCATGCCTC CTTCCCAGGA   660
 Q  R  G   A  F  H   P  S  N  *                                    205

                                  <-----#1484
CCCAGGTCTG CTGCCCAGGA ACCCTCCAGA CCTCCAAGTG AAGAATAAAG TGTGGCTCGT   720

G                                                                   721
```

Fig. 2. Nucleotide sequence (GenBank accession No. L10035) of human β*B2-cry* cDNA and deduced aa sequence. There are 12 bp of 5′*UTR* and 94 bp of 3′*UTR*. The boxed sequences were used to make specific oligos used for the PCR amplification reactions.

EXPERIMENTAL AND DISCUSSION

(a) PCR amplification of cDNAs

The sequence information from mouse genomic clones obtained by this laboratory was important in generating a 30-mer oligo primer corresponding to the first exon of the β*B2-cry* gene. This primer was used for PCR amplification using RNA isolated from human lens tissue along with a 3′ oligo based on mβ*B2-cry* cDNA sequences. The agarose gel electrophoresis of the amplified product showed a band of approximately 750 bp (Fig. 1). Several PCR clones were sequenced using as primers oligos based on mβ*B2-cry* cDNA sequences. Fig. 2 shows the nt sequence of the hβ*B2-cry* cDNA (GenBank accession No. L10035) and the encoded polypeptide of 205 aa. The aa coding sequence begins at nt 13 and extends through nt 667, leaving a 12-nt 5′*UTR* and a 94-nt 3′*UTR*.

(b) Construction of λgt11 cDNA libraries

To insure no PCR artifacts were introduced in the sequence, cDNA libraries were also made from human adult and fetal lens tissue RNAs and cloned into the *Eco*RI site of λgt11 vector. Both cDNAs were first

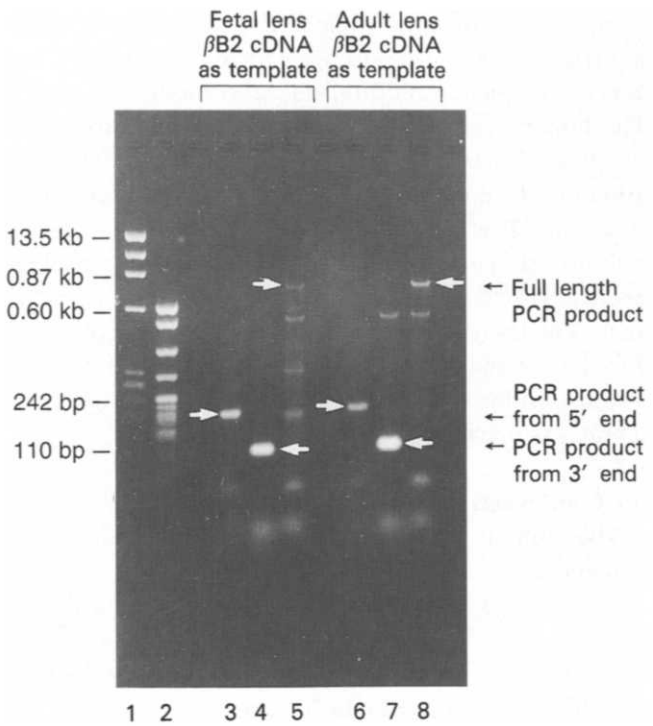

Fig. 3. Agarose gel of PCR-generated fragments from hB*B2-cry* cDNAs. Lanes 1 and 2 are DNA markers (ϕX174 RF DNA/*Hae*III and pBR322 DNA/*Msp*I, respectively). Lanes 3, 4, 5 contain PCR products using fetal lens RNA and lanes 6, 7, 8 contain PCR products using adult lens RNA. Lanes 3 and 6 show a 192 bp PCR fragment (indicated by white arrows located to the left of the bands) that contains exons 1, 2 and 3 (amplifying a portion of the 5′ coding region of the cDNA). Lanes 4 and 7 show a 108 bp PCR fragment (indicated by white arrows located to the right of the bands) that contains part of exon 6 and 3′*UTR* (amplifying a portion the of 3′ region of the cDNA). Lanes 5 and 8 show multiple sized PCR fragments of the same PCR product, with the uppermost band (indicated by the white arrows near each band) representing the size of the full-length cDNA that contains exons 1 to 6 plus 3′*UTR*. **Methods:** Total RNA was extracted from lenses obtained from several donor eyes using RNAzol B (Tel-Test, Inc., Friendswood, TX, USA). The ages of the individual subjects ranged from 11 to 47 years old. Also RNA was extracted from human fetal lenses that ranged from 20 to 24 weeks gestation. cDNA was synthesized from the total RNA preps using the cDNA synthesis (Pharmacia LKB, Piscataway, NJ, USA). Depending on the reaction conditions there may be a decrease in extension rate or processivity of *Taq* DNA polymerase which may result in shorter PCR products being generated. In this case the variable that may have caused this to occur was the primer concentration. The primer concentration may have contributed to the additional bands seen at 300 and 600 bp since these bands, when probed with mβ*B2-cry* cDNA, hybridized strongly under stringent conditions (data not shown). Also the formation of primer-dimers was seen which is also attributed to primer concentration. Sequence information obtained from mouse clones and sequence information from human clones (Brakenhoff et al., 1992) allowed positioning of exons. The oligos used in the PCR amplification of the cDNAs are boxed in Fig. 2. Oligo #1482 and the 5′ oligo listed in Fig. 1 (without the *Eco*RI-*Not*I adapter) were used to check the 5′ region and oligos #1483 and #1484 were used to check the 3′ region of the generated cDNAs. To check for the presence of the full-length sequence, the 5′ oligo listed in Fig. 1 (without the *Eco*RI-*Not*I adapter) and #1484 were used. All oligos used as primers in experiments were based on sequence information obtained from mouse β*B2-cry* cDNA.

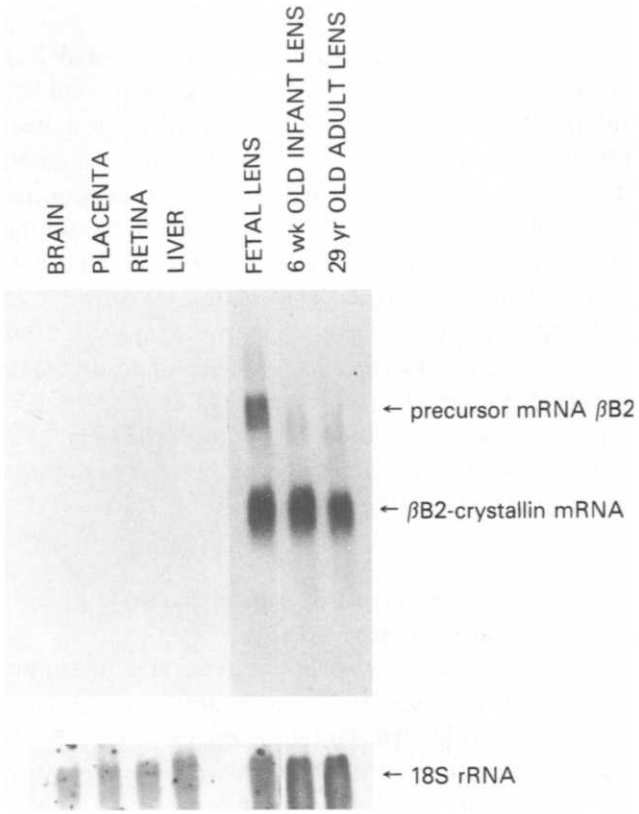

Fig. 4. Northern blot analysis of total RNAs (human brain, placenta, retina, and liver (Clontech, Palo Alto, CA, USA), and total RNAs isolated from human adult and fetal lens tissues. Total RNAs were isolated as described in Fig. 3. **Methods:** 10 μg of total RNA was loaded onto 1% agarose-2.2 M formaldehyde gel, separated by electrophoresis, blotted onto positively charged nylon membrane (Boehringer Mannheim), and fixed by UV irradiation (Stratagene UV Stratalinker). The probes were a full-length mβ*B2-cry* cDNA nonradioactively labelled with digoxigenin-UTP using the Genius DNA labeling kit and an 18S rRNA cDNA probe also labelled using the Genius kit.

checked for the presence of the full-length β*B2-cry* cDNA by PCR. Using a set of oligos (boxed sequences in Fig. 2, as well as, the 5′ oligo mentioned in Fig. 1), it was determined that both sets of cDNAs contained the full-length sequence (Fig. 3). The adult λgt11 library was chosen for screening the human β*B2-cry* cDNA sequences using the mβ*B2-cry* cDNA as a probe (Chambers and Russell, 1991). Several positive clones were sequenced. Both strands of each clone were sequenced. Since almost all of the hβ*B2-cry* cDNA sequence was generated by using oligos as primers that were based on the mβ*B2-cry* cDNA sequence, the sequencing strategy was similar to the one used in Chambers and Russell, 1991. Only four additional oligos (based on sequences generated from the human λgt11 β*B2-cry* cDNA clones) were needed to complete the sequencing of the hβ*B2-cry* cDNA. No differences in the sequences of the λgt11 cDNA clones and those generated from the PCR cloning were seen, indicating that the PCR-generated sequence contained no artifacts.

(c) Analysis of mRNA expression

Northern blot analysis of total RNAs isolated from human adult lens, retina, liver, placenta, brain, and fetal lens tissues is shown in Fig. 4. A transcript of approx. 0.76 kb was found for the adult and fetal lens tissues only. These data would suggest that βB2-*cry* expression may be lens specific (Fig. 4). It has been shown by others that β-*cry* genes exhibit tissue specific expression and are developmentally and spatially regulated (Wistow et al., 1988). Also seen was a larger transcript of approx. 2.0 kb that may represent a precursor intermediate or unprocessed mRNA for βB2-Cry. The size of the βB2-*cry* gene in humans has been shown to be approximately 28 kb (Aarts et al., 1987). Therefore, this 2.0-kb band could represent an intermediate RNA.

(d) Comparison of mouse βB2 with human βB2 sequences; deduced aa sequences

Sequence comparison of the complete cDNAs showed that the human nt sequence was slightly more similar to the bovine than to the murine or rat sequences (88.5% identity; 87.2% identity; and 83.6% identity, respectively). The sequence similarities among the species are summarized in Table I. If one compares only the sequences that code for domains 1 and 2 and not the sequences that code for the N and C termini, the identity increases to 91% between human and bovine, to 88.8% between human and mouse, and to 84.4% between human and rat.

The sequence comparison of the deduced aa sequences is also summarized in Table I. The aa sequences of the human, bovine and mouse were identical in motif 2 and almost complete similarity is seen in motifs 1, 3, and 4. The human sequence has the same 3′ phosphorylation site that is present in the bovine and rat proteins but absent in the mouse. The role of phosphorylation is controversial. This modification may be important for the stability of the polypeptide or may influence protein-protein associations that are essential for lens transparency. The length of the mouse βB2-Cry polypeptide is a 2 aa longer than both the human and bovine proteins. The reason for the shorter polypeptide in the higher vertebrates is unclear.

(e) Conclusions

The human β*B2-cry* cDNA has been cloned and sequenced.

(*1*) The full length cDNA is 721 bp, with a 205 aa ORF.

(*2*) The hβ*B2-cry* shows slightly greater similarity to the bβ*B2-cry* than to the mβ*B2-cry* or rβ*B2-cry*.

(*3*) The human βB2-Cry protein is two aa shorter than that of the mouse.

(*4*) Northern blot analysis suggests that β*B2-cry* is lens specific.

The gene is expressed in 20–24 week fetal lens tissue.

The availability of the full-length hβ*B2-cry* cDNA provides a useful tool to begin to investigate defects or changes in the human βB2-Cry protein which may be associated with congenital cataract formation.

TABLE I

Comparison of hβB2, bβB2, mβB2, rβB2-*cry* nt and deduced aa sequence similarity[a]

	Sequence similarity			
	nt sequence		aa sequence	
	% Identity	% Similarity	% Identity	% Similarity
Full-length cDNA:				
Human-bovine	88.5	88.5	90.3	93.0
Human-mouse	87.2	87.2	89.7	90.1
Human-rat	83.6	83.6	82.9	88.0
Coding regions corresponding to motifs 1–4 (exons 3–6):				
Human-bovine	91.09	91.09	97.75	99.43
Human-mouse	88.82	88.82	97.75	98.87
Human-rat	84.47	84.47	92.13	93.82

[a]The nt and aa sequences of human (hβB2), bovine (bβB2), mouse (mβB2), and rat (rβB2) *cry* were analyzed and compared by using the GCG software version 7.0. % Identity refers to aa sequences which were completely conserved. % Similarity refers to aa sequences which included conservative aa changes (Devereux et al., 1984).

ACKNOWLEDGEMENTS

We wish to thank the National Disease Research Interchange (NDRI), Philadelphia, PA, for supplying the donor material.

Note: Since submission of this manuscript, the aa sequence of the human lens βB2-Cry protein has been published and the deduced aa sequence reported in this paper matches 100% with the sequence that was determined by fast atom bombardment mass spectrometry (Miesbauer et al., 1993).

REFERENCES

Aarts, H.J.M., Den Dunnen, J.T., Lubsen, N.H. and Schoenmakers, J.G.G.: Linkage between the βB2 and βB3 crystallin genes in man and rat: a remnant of an ancient β-crystallin gene cluster. Gene 59 (1987) 127–135.

Aarts, H.J.M., Lubsen, N.H. and Schoenmakers, J.G.G.: Crystallin gene expression during rat lens development. Eur. J. Biochem. 183 (1989a) 31–36.

Aarts, H.J.M., Jacobs, E.H., Van Willigen, G., Lubsen, N.H. and Schoenmakers, J.G.G.: Different evolution rates within the lens-specific β-crystallin gene family. J. Mol. Evol. 28 (1989b) 313–321.

Benedek, G.B.: Theory of transparency of the eye. Appl. Opt. 10 (1971) 459–473.

Bettelheim, F.A. and Siew, E.L.: Biological-physical basis of lens transparency. In: McDevitt, D.S. (Ed.), Cell Biology of the Eye. Academic Press, New York, 1982, pp. 243–297.

Brakenhoff, R.H., Aarts, H.J.M., Schuren, F., Lubsen, N.H. and Schoenmakers, J.G.G.: The second human βB2-crystallin gene is a pseudogene. Exp. Eye Res. 54 (1992) 803–806.

Chambers, C. and Russell, P.: Deletion mutation in an eye lens β-crystallin. J. Biol. Chem. 266 (1991) 6742–6746.

Delaye, M. and Tardieu, A.: Short range order of crystallin proteins accounts for eye lens transparency. Nature 302 (1983) 415–417.

Den Dunnen, J.T., Moormann, R.J. and Schoenmakers, J.G.G.: Rat lens β-crystallins are internally duplicated and homologous to γ-crystallins. Biochim. Biophys. Acta 824 (1985) 295–303.

Devereux, J., Haeberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12 (1984) 387–395.

Hejtmancik, J.F., Thompson, M.A., Wistow, G.J. and Piatigorsky, J.: cDNA and deduced protein sequence for the βB1-crystallin polypeptide of the chicken lens: conservation of the PAPA sequence. J. Biol. Chem. 261 (1986) 982–987.

Hogg, D., Gorin, M., Heinzmann, C., Zollman, S., Mohandas, T., Klisak, I., Sparkes, R., Breitman, M.L., Tsui, Lap-Chee and Horowitz, J.: Nucleotide sequence for the cDNA of the bovine βB2-crystallin and assignment of the orthologous human locus to chromosome 22. Current Eye Res. 6 (1987) 1335–1342.

Inana, G., Shinohara, T., Maizel, J.V. and Piatigorsky, J.: Evolution and diversity of the crystallins. Nucleotide sequence of a β-crystallin mRNA from the mouse lens. J. Biol. Chem. 257 (1982) 9064–9071.

Lubsen, N.H., Renwick, J.H., Tsui, L.-C., Breitman, M.L. and Schoenmakers, J.G.G.: A locus for a human hereditary cataract is closely linked to the γ-crystallin gene family. Proc. Natl. Acad. Sci. USA 84 (1987) 489–492.

Miesbauer, L.R., Smith, J.B. and Smith, D.L.: Amino acid sequence of human lens βB2-crystallin. Prot. Sci. 2 (1993) 290–291.

Peterson, C.A. and Piatigorsky, J.: Preferential conservation of the globular domains of the βA3/A1-crystallin polypeptide of the chicken eye lens. Gene 45 (1986) 139–147.

Piatigorsky, J.: Lens crystallins and their gene families. Cell 38 (1984) 620–621.

Quax-Jeuken, Y., Janssen, C., Quax, W., Van den Heuvel, R. and Bloemendal, H.: Bovine β-crystallin complementary DNA clones: alternating proline/alanine sequence of βB1 subunit originates from a repetitive DNA sequence. J. Mol. Biol. 180 (1984) 457–472.

Rodriguez, I.R., Gonzalez, P., Zigler Jr., J.S. and Borras, T.: A guinea-pig hereditary cataract contains a splice-site deletion in a crystallin gene. Biochim. Biophys. Acta 1180 (1992) 44–52.

Van Rens, G.L.M., Driessen, H.P.G., Nalini, V., Slingsby, C., De Jong, W.W. and Bloemendal, H.: Isolation and characterization of cDNAs encoding βA2- and βA4-crystallins: heterologous interactions in the predicted βA4-βB2 heterodimer. Gene 102 (1991) 179–188.

Wistow, G. and Piatigorsky, J.: Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. Annu. Rev. Biochem. 57 (1988) 479–504.

Zigler Jr., J.S.: Animal models for the study of maturity onset and hereditary cataract. Exp. Eye Res. 50 (1990) 651–657.