Vol. 146, No. 1, 1987 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1987 Pages 332-338

cDNA CLONES ENCODING BOVINE γ-CRYSTALLINS

Regine E. Hay, Wendell D. Woods, Robert L. Church, and J. Mark Petrash

Department of Ophthalmology,
Emory University School of Medicine,
Atlanta, GA 30322

Received June 1, 1987

We have determined the nucleotide sequence of two bovine lens γ-crystallin cDNA clones, pBLγII-1 and pBLγIII-1. The 644 bp cDNA insert of pBLγII-1 contains coding information for the entire amino acid sequence of bovine γII-crystallin. The 497 bp cDNA insert of pBLγIII-1 encodes a homologous but different γ-crystallin polypeptide, and appears to lack the coding information for the C-terminal 17 amino acid residues. While the nucleotide and predicted amino acid sequences of the coding regions of the clones show a high degree of homology, the untranslated leader sequences are relatively dissimilar. The leader sequence of pBLγIII-1 is strikingly homologous to a portion of a rabbit immunoglobulin α-heavy chain mRNA.

The ocular lens contains a high concentration of proteins, called crystallins, which account for approximately 90% of the total soluble protein in this tissue (1). In mammals, the crystallins may be divided into three immunologically distinct classes, α, β, and γ, each of which is composed of multiple proteins with related primary structure. The β- and γ-crystallins appear to be evolutionarily related (2) and are each encoded by multigene families (3,4). The γ-crystallins are composed of at least 5 homologous polypeptides in bovine (5) and rodent (6) lenses. Transcription and synthesis of γ-crystallins, which are restricted to the differentiating cortical fiber cells of the lens, are also developmentally regulated (7,8). A pattern of gene expression which is controlled at precise differentiation and developmental stages implies that a unique complement of γ-crystallins may be required for normal lens function. How these gene products may interact among themselves and/or with other structural components of the lens is not understood at present. In order to delineate the structural and functional relationships among the γ-crystallins, we have initiated studies to isolate their correspond-

0006-291X/87 $1.50

ing genes. In this work, we report the isolation and characterization of two bovine γ-crystallin cDNAs.

MATERIALS AND METHODS

Materials. All enzymes were purchased from Bethesda Research Laboratories. Isotopes were from New England Nuclear. All oligodeoxynucleotide sequencing primers were prepared with an Applied Biosystems Model 381A DNA synthesizer using reagents purchased from the manufacturer. The MicroGenie DNA sequence data base (Version 3.2, updated Jan/Feb 1985) was from Beckman Instruments.

Isolation and Nucleotide Sequencing of cDNA Clones. Plasmid clones containing γ-crystallin cDNAs were isolated from a total bovine lens cDNA library as described previously (9). Restriction fragments from the cDNA inserts were cloned into M13 mp18 and mp19 either directly or after subcloning first into pUC18 or pUC19. Nucleotide sequencing was carried out by the dideoxy termination method (10) using either universal sequencing primers or synthetic primers constructed to anneal to internal cDNA sequences. Sequencing data were compiled using the NUMSEQ analysis program (11) contained in a software package made available by David Mount, University of Arizona, Tucson.

RESULTS AND DISCUSSION

Isolation and Characterization of cDNA Clones. A total of 8 independent γ-crystallin cDNA clones were selected from a bovine lens cDNA library by colony hybridization screening with a synthetic 51-mer oligonucleotide probe (9). Total cDNA inserts released from these clones ranged in size from 350-725 bp, and all cDNA inserts contained an internal Pst-I recognition site. Additional restriction mapping and partial nucleotide sequencing of most of the cDNA inserts showed that the clones could be assigned to one of two classes, designated γII and γIII. In the present study, we have analyzed the nucleotide and predicted amino acid sequences of pBLγII-1 and pBLγIII-1, cDNA clones representative of each class.

Nucleotide and Amino Acid Sequence Comparisons. The nucleotide sequences of pBLγII-1 and pBLγIII-1 are aligned for comparison in Figure 1. Excluding the poly(A) tail and dG-dC homopolymer tracts, pBLγII contains a cDNA insert of 624 base pairs (bp) in length. An open reading frame is found in pBLγII commencing with the methionine codon at base number 32. The sequences flanking this ATG (MET) initiation codon conform to a consensus sequence (PuXXATGG) flanking most eucaryotic translational initiation sites (12). Translation of the remaining sequences predicts a polypeptide containing 174 amino acid residues. The 66 bp nontranslated region located downstream from the termination codon contains a second in-frame termination codon located at position 587 as well as a poly-

```
                       10            20           30           40            50           60           70
pBLγII      C AGC GAA CAC CTC CTG CTG CCT TTC GCC AAC ATG GGG AAG ATC ACT TTT TAC GAG GAC CGG GGC TTC CAG GGC CAC
                      ***  ** ***  *   *  ***   * **                    *   *                                 *
pBLγIII             A ACT CCA TCC CCG CGT CCA GCA GCC ATG GGG AAG ATC ACC TTC TAC GAG GAC CGG GGC TTC CAG GGC CGC

BLγII                                                 MET Gly Lys Ile Thr Phe Tyr Glu Asp Arg Gly Phe Gln Gly His
BLγIII                                                MET Gly Lys Ile Thr Phe Tyr Glu Asp Arg Gly Phe Gln Gly Arg
                                                           1                                   10

                 80           90           100           110          120          130           140          150
pBLγII        TGC TAC GAG TGC AGC AGC GAC TGC CCC AAC CTG CAG CCC TAT TTC AGC CGC TGT AAC TCC ATC CGC GTG GAC AGC
              **    *                     **  *                 *   * * * *         *   *   * * *   *
pBLγIII       CAC TAT GAG TGC AGC AGC GAC CAC TCC AAC CTG CAG CCT TAC CTG GGC CGC TGC AAT TCG GTG CGT GTG GAC AGC

BLγII         Cys Tyr Glu Cys Ser Ser Asp Cys Pro Asn Leu Gln Pro Tyr Phe Ser Arg Cys Asn Ser Ile Arg Val Asp Ser
BLγIII        His Tyr Glu Cys Ser Ser Asp His Ser Asn Leu Gln Pro Tyr Leu Gly Arg Cys Asn Ser Val Arg Val Asp Ser
                                   20                                      30

                       160           170          180          190           200          210          220
pBLγII        GGC TGC TGG ATG CTG TAT GAG CGC CCC AAC TAC CAG GGC CAC CAG TAC TTC CTG CGG CGC GGC GAC TAC CCC GAC
                *             * *   *    **           *  *      *                             *       *
pBLγIII       GGT TGC TGG ATG ATC TAC GAG CAG CCC AAC TAT CTG GGC CCC CAG TAC TTC CTG CGG CGC GGT GAC TAT CCC GAC

BLγII         Gly Cys Trp Met Leu Tyr Glu Arg Pro Asn Tyr Gln Gly His Gln Tyr Phe Leu Arg Arg Gly Asp Tyr Pro Asp
BLγIII        Gly Cys Trp Met Ile Tyr Glu Gln Pro Asn Tyr Leu Gly Pro Gln Tyr Phe Leu Arg Arg Gly Asp Tyr Pro Asp
               40                                      50                                      60

                 230          240          250           260          270          280           290          300
pBLγII        TAC CAG CAG TGG ATG GGC TTC AAC GAC TCC ATC CGC TCC TGC CGC CTC ATC CCG CAA CAC ACC GGC ACT TTC AGA
                                      *                                         *           * *     *   **    *
pBLγIII       TAC CAG CAG TGG ATG GGC CTC AAC GAC TCC ATC CGC TCC TGC CGC CTC ATC CCC --- CAC GCT GGC TCT CAC AGG

BLγII         Tyr Gln Gln Trp Met Gly Phe Asn Asp Ser Ile Arg Ser Cys Arg Leu Ile Pro Gln His Thr Gly Thr Phe Arg
BLγIII        Tyr Gln Gln Trp Met Gly Leu Asn Asp Ser Ile Arg Ser Cys Arg Leu Ile Pro --- His Ala Gly Ser His Arg
                                   70                                      80

                       310           320          330          340           350          360          370
pBLγII        ATG AGA ATC TAT GAG AGA GAT GAC TTC AGA GGA CAG ATG TCA GAG ATC ACA GAC GAT TGT CCC TCT CTT CAA GAC
              * *     * *           *   *   *  *      *         **          *   *   *   * *
pBLγIII       CTC AGA CTT TAT GAG AGG GAA GAT TAC AGA GGC CAG ATG ATA GAG ATC ACT GAA GAC TGC TCC TCT CTT CAA GAC

BLγII         Met Arg Ile Tyr Glu Arg Asp Asp Phe Arg Gly Gln Met Ser Glu Ile Thr Asp Asp Cys Pro Ser Leu Gln Asp
BLγIII        Leu Arg Leu Tyr Glu Arg Glu Asp Tyr Arg Gly Gln Met Ile Glu Ile Thr Glu Asp Cys Ser Ser Leu Gln Asp
               90                                     100                                     110

                 380          390          400           410          420          430           440          450
pBLγII        CGC TTC CAC CTC ACT GAG GTT CAC TCC CTC AAC GTG CTG GAG GGT TCC TGG GTC CTC TAT GAG ATG CCA AGC TAC
                          *           *         *       *           *   *                         *     *  *
pBLγIII       CGC TTC CAC TTC ACT GAG ATT CAC TCT CTC AAT GTG CTG GAA GGC TCC TGG GTC CTC TAT GAG TTG CCC AAC TAC

BLγII         Arg Phe His Leu Thr Glu Val His Ser Leu Asn Val Leu Glu Gly Ser Trp Val Leu Tyr Glu Met Pro Ser Tyr
BLγIII        Arg Phe His Phe Thr Glu Ile His Ser Leu Asn Val Leu Glu Gly Ser Trp Val Leu Tyr Glu Leu Pro Asn Tyr
                                  120                                     130

                       460           470          480          490           500          510          520
pBLγII        AGG GGA AGG CAG TAC CTG CTG AGG CCA GGG GAG TAC AGG AGA TAT CTT GAC TGG GGG GCA ATG AAT GCC AAA GTT
              *     * *       *         *                       * *   *  *
pBLγIII       CGG GGG CGG CAG TAT CTG CTG CGG CCA GGG GAG TAC AGG CGC TAC CAT GAC TGG G

BLγII         Arg Gly Arg Gln Tyr Leu Leu Arg Pro Gly Glu Tyr Arg Arg Tyr Leu Asp Trp Gly Ala Met Asn Ala Lys Val
BLγIII        Arg Gly Arg Gln Tyr Leu Leu Arg Pro Gly Glu Tyr Arg Arg Tyr His Asp Trp
              140                                     150                                     160

                530          540          550           560          570          580           590          600
pBLγII       GGT TCT TTA AGA CGG GTG ATG GAT TTT TAT TGA AGT TTT AAA TTC CCC ACT TTT CTC CTT TAG AAT CTA ATA AAA
BLγII        Gly Ser Leu Arg Arg Val Met Asp Phe Tyr TER                                 TER
                                 170

                      610           620
pBLγII       GAT TTA GCC TTG TGT TTC TGG CCA_20
```

Figure 1. Nucleotide and predicted amino acid sequences of two bovine lens γ-crystallin cDNAs, pBLγII-1 and pBLγIII-1. For comparison, the nucleotide sequences from pBLγIII are aligned and numbered relative to that of pBLγII. Their predicted amino acid sequences (BLγII and BLγIII) are aligned accordingly. Nucleotide differences between the cDNAs are indicated by "*". A gap (---) has been introduced into the BLγIII sequence at residue #83 in order to maximize alignment with BLγII. Both cDNAs contained poly(dG) and (dC) tails of approximately 20 nucleotides at either end of the cloned inserts (not shown). Stop codons in pBLγII are translated as TER. The 20 nucleotide poly (A) tail in pBLγII is shown as A_{20}.

adenylation signal (AATAAA) located 30 bp upstream from the poly(A) tail. The nucleotide sequence of pBLγII-1 is almost identical to that of pγCRYB1, a bovine γ-crystallin cDNA clone described previously (13). A total of 7 nucleotide differences are found between pBLγII-1 and pγCRYB1. The only amino acid difference predicted from these cDNAs is at threonine-119, which was previously reported as serine in pγCRYB1 (13). The predicted amino acid sequence from pBLγII-1, including threonine at residue 119, is consistent with X-ray crystallographic analysis of calf γII-crystallin (14; Christine Slingsby, personal communication).

The γ-crystallin cDNA clone pBLγIII-1 contains 497 bp of cDNA insert. Similar to pBLγII-1, the cDNA insert of pBLγIII-1 begins 25 bp upstream from a methionine initiation codon at base number 26. The ensuing coding sequence may be translated to yield a polypeptide containing 155 amino acid residues. Therefore, it appears that pBLγIII-1 is missing the sequences encoding the last 17 amino acids of the final gene product.

The overall nucleotide and amino acid sequence homology between the two γ-crystallins is approximately 82% and 83%, respectively. The amino acid sequence encoded by pBLγIII-1 is notable in the absence of Gln-83, an amino acid residue located in the connecting peptide region of the γ-crystallin molecule. Most other amino acid differences predicted by these cDNAs reflect structurally conservative changes (e.g. Leu 44-Ile, Arg 47-Gln, Asp 107-Glu) and would probably have little effect on the tertiary structure. Other amino acids thought to play a role in stabilizing the tertiary structure of the γ-crystallin molecule (14) are all well conserved (e.g. Tyr-6, Phe-56, Trp-68). Comparison of the amino acid sequences corresponding to each of the structural motifs revealed that these γ-sequences differ the most in motif 3 (75.6% homology), whereas the sequences corresponding to motifs 1,2, and 4 are more highly conserved (>82% homology) (Table 1). This is similar to that observed with the mouse (3), rat (6), and human (15) γ-crystallin genes.

The most striking nucleotide sequence divergence between pBLγII-1 and pBLγIII-1 occurs in the untranslated sequences located immediately upstream from their respective initiation codons. These sequences are only 36% homologous when compared without alignment corrections, in contrast to their

Table 1. The nucleotide and predicted amino acid sequences from PBγII-1 and pBLγIII-1 were aligned as shown in Figure 1. Boundaries of the structural motifs were taken from (14)

SEQUENCE HOMOLOGIES BETWEEN pBLγII-1 AND pBLγIII-1

	Nucleotide	Amino Acid
Leader	36.0	-
Motif 1	83.3	82.5
Motif 2	88.6	85.1
Motif 3	80.5	75.6
Motif 4*	85.2	89.7
Overall	82.1	82.7

"*" Comparison of sequences in motif 4 was limited by the incomplete coding information contained in pBLγIII-1.

85% homologous coding sequences. While it is possible that these leader sequences may have evolved without selective pressure, it is also possible that they may be involved in modulating the translational efficiency of their respective mRNAs. In order to identify other gene sequences of known functional significance which may be homologous to the leader sequences of pBLγII-1 and pBLγIII-1, a homology search with a gene bank data base was conducted. Unexpectedly, the pBLγIII-1 leader sequence showed strong homology (20 matches out of 24 nucleotides) with the sequence of a rabbit IgA α-heavy chain mRNA (16). The corresponding α-chain mRNA sequence codes for amino acid residues 398-405 in the CH3 domain of the immunoglobulin molecule. The functional significance of this sequence homology is not presently understood. No significant homologies with other gene bank sequences were observed with either of the leader sequences.

Sequence comparisons with γ-crystallins from other species. We have attempted to correlate the nucleotide and amino acid sequences of pBLγII-1 and pBLγIII-1 with γ-crystallin gene sequences from other organisms. The rat γ1-2 and mouse γ3 genes show the greatest homology with our pBLγII-1. None of the

Table 2. The predicted amino acid sequences from pBLγII-1 and pBLγIII-1 were aligned with the corresponding sequences predicted from rat (6), mouse (3), and human (15) γ-crystallin genes

	AMINO ACID SEQUENCE HOMOLOGIES	
	pBLγII-1	pBLγIII-1*
Rat γ1-1	81	79
γ1-2	91	81
γ2-1	81	80
γ2-2	82	89
γ3-1	81	84
γ4-1	79	83
Mouse γ1	79	86
γ2	81	84
γ3**	85	65
γ4	81	77
Human G1ψ	64	78
G2ψ	74	84
G3	76	76
G4	76	87

"*" In the case of pBLγIII-1, alignments were made with amino acid residues 1-157.

"**" Comparison with mouse γ3 sequences was limited to residues 37-173.

individual human γ-crystallin genes were distinctly homologous with pBLγII-1 (Table 2). Apart from γII, the amino acid sequences for the bovine γ-crystallins are not known, precluding our assignment of pBLγIII-1 to a specific bovine γ-crystallin fraction. However, it appears that pBLγIII-1 is orthologous to rat γ2-2, mouse γ1, and human G4 genes (Table 2).

It has been postulated that amino acid sequence differences among the various γ-crystallins may be responsible for the unique biophysical characteristics of each γ component (14). Through site-directed mutagenesis of selected amino acid residues contained in pBLγII-1, and subsequent expression of the altered sequences in cultured cells, various aspects of cryoprecipitation behavior (17) and spectroscopic properties (18) of the γ-crystallins can be studied.

ACKNOWLEDGEMENTS This work was supported by research grants from the National Eye Institute (R23 EY05639 to JMP, R01 EY05877 and 1C06 EY06307 to