Vol. 191, No. 3, 1993
March 31, 1993
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 902-907

MOLECULAR CLONING AND SEQUENCING OF ZETA-CRYSTALLIN/QUINONE REDUCTASE cDNA FROM HUMAN LIVER

Pedro Gonzalez, P. Vasantha Rao and J. Samuel Zigler, Jr.

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

Received January 22, 1993

Zeta-crystallin is an enzyme-crystallin highly expressed in the lens of some hystricomorph rodents and camels. It has been shown to have a novel NADPH: quinone oxidoreductase activity and is present at enzymatic levels in a variety of tissues from various mammals. We report here the cDNA cloning of ζ-crystallin from a human liver library. One clone with the complete open reading frame was obtained. Ten nucleotides of the 5' and 796 of the 3' non-translated regions are present in the clone including two possible polyadenylation signals. The deduced amino acid sequence is 328 residues long with a calculated molecular mass of 34910 daltons and isoelectric point of 8.73. It shows 84% identity with the guinea pig protein.

Zeta-crystallin is a major protein in the lens of some hystricomorph rodents (1) and camels (2). Like other enzyme-crystallins it is also present at enzymatic levels in certain non-lenticular tissues of different species (3). ζ-Crystallin has been shown to possess an NADPH dependent quinone reductase activity which appears to be novel, because unlike other one electron transfer quinone reductases, it has a limited substrate specificity, it is not membrane bound, and does not contain a flavin moiety (4).

The complete protein sequence deduced from cDNA clones in guinea pig (5) contained 328 amino acids and showed a significant relationship to the enzymes of the long-chain, zinc-containing alcohol dehydrogenase (ADH) family (6). Studies on the ζ-crystallin gene from guinea pig yielded two findings of particular interest. First, a genomic mutation causing the deletion of 34 amino acids of the ζ-protein was clearly associated with a hereditary nuclear cataract in a line of strain 13 guinea pigs (7). This is the only case described so far of a cataract due to a mutation in an enzyme-crystallin. Secondly, ζ-crystallin of guinea pig is the first enzyme-crystallin gene with two distinct promoters, one responsible for high expression in the lens and the other for expression at enzymatic level in other tissues (*). Modification in the pattern of expression of a gene leading to the acquisition of a new

* P. Gonzalez et al., manuscript in preparation.

0006-291X/93 $4.00

function without gene duplication has been proposed as an evolutionary mechanism to explain the recruitment of enzymes as lens proteins (8). The presence of a second "lens" promoter represents a possible means of such "gene sharing". Establishment of such a mechanism for ζ-crystallin requires analysis of the gene in species such as guinea pig where ζ is a crystallin and also in species where ζ is not expressed at high level in the lens.

In this study we report the cloning and sequencing of ζ-crystallin/quinone reductase cDNA from human liver, and the deduced amino acid sequence of the protein. This study provides the first data on ζ-crystallin from a species in which it exists only in very low amounts in the lens. Its presence suggests that this protein has a distinct physiological function in lens apart from its presumed role as a structural lens protein (crystallin) in those species where it is expressed at high levels. The present study is also a step toward analysis of the structure of the human ζ-gene and elucidation of the mechanism responsible for its recruitment as a crystallin in certain species.

MATERIALS AND METHODS

Two λgt10, oligo dT primed libraries commercially available were used for this study: human fetal liver cDNA library (Clontech cat. HL1064), and adult human liver cDNA library (Clontech cat. HL1115a). The probe used for the screening was a fragment of the guinea pig ζ-crystallin cDNA (nucleotides 147 to 690 from the published sequence, Rodokanaki et al., 1989) obtained by the polymerase chain reaction and labeled by the random primer method with (^{32}P)dCTP. About 10^6 plaques were screened using nitrocellulose filters (S&SNC Schleicher & Schuell). Hybridizations were performed at 60°C in 3X SSC, 10% SDS, 0.5% Denhardt's and 100 μg/ml salmon sperm DNA (Sigma) with washes at 60°C in 2X SSC, 0.1% SDS.

The insert of the positive clone was released with *Eco*RI and subcloned in the vector BlueScript KS II+ (Stratagene) previously linearized and dephosphorylated. After overnight ligation at 16°C, DNA was used to transform DH5α competent cells (BRL). The recombinant plasmids were purified with Qiagen columns and both strands were sequenced by the dideoxy chain termination method (9) with (^{35}S)dATP and Sequenase II (USB).

Computer analysis was performed using the GCG package of programs in the NIH UNIX computer.

RESULTS AND DISCUSSION

No positive clones were found as the result of screening approximately 10^6 plaques from the human fetal liver library. After screening a similar number of plaques from the adult liver library, three positive clones: HL101, HL102, and HL103 were isolated. HL101 and HL102 were sibling clones which were discarded when analysis of the sequence revealed some recombination with other genes not related to ζ-crystallin. The sequence of the 1796 nucleotide insert in clone HL103 revealed no evidence of recombination, and

this clone was selected for a complete sequence analysis. The small number of positive clones obtained in a relatively large number of plaques screened probably reflects a low level of expression of ζ-crystallin in human liver.

The putative initiation codon ATG (Met) was at nucleotide 11-13 (numbered after the linkers) and the sequence surrounding this ATG codon (ATCACCATGG) matches the consensus for the translation initiation region of vertebrate genes (10). The termination codon TGA was located at positions

```
                A  T  G  Q  K  L  M  R  A  V  R  V  F  E  F  G  G  P  E  V  L  K    22
  CTAGATCACCATGGCGACTGGACAGAAGTTGATGAGAGCTGTTAGAGTTTTTGAATTTGGTGGGCCAGAAGTCCTGAAA   79

 L  R  S  D  I  A  V  P  I  P  K  D  H  Q  V  L  I  K  V  H  A  C  G  V  N  P  V    49
TTGCGATCAGATATTGCAGTACCGATTCCAAAAGACCATCAGGTTCTAATCAAGGTCCATGCATGTGGTGTCAACCCCGTG  160

 E  T  Y  I  R  S  G  T  Y  S  R  K  P  L  L  P  Y  T  P  G  S  D  V  A  G  V  I    76
GAGACATACATTCGCTCTGGTACTTATAGTAGAAAACCACTCTTACCCTATACTCCTGGCTCAGATGTGGCTGGGGTGATA  241

 E  A  V  G  D  N  A  S  A  F  K  K  G  D  R  V  F  T  S  S  T  I  S  G  G  Y  A   103
GAAGCTGTTGGAGATAATGCATCTGCTTTCAAGAAAGGTGACAGAGTTTTCACTAGCAGCACGATCTCTGGGGGTTATGCA  322

 E  Y  A  L  A  A  D  H  T  V  Y  K  L  P  E  K  L  D  F  K  Q  G  A  A  I  G  I   130
GAGTATGCTCTTGCAGCAGACCACACTGTTTACAAACTACCTGAAAAACTGGACTTTAAACAAGGAGCTGCCATCGGCATT  403

 P  Y  F  T  A  Y  R  A  L  T  H  S  A  C  V  K  A  G  E  S  V  L  V  H  G  A  S   157
CCATATTTTACTGCTTATCGAGCTCTGATCCACAGTGCCTGTGTGAAAGCTGGAGAGAGTGTTCTGGTTCATGGGGCAAGT  484

 G  G  V  G  L  A  A  C  Q  I  A  R  A  Y  G  L  K  I  L  G  T  A  G  T  E  E  G   184
GGAGGAGTTGGATTAGCAGCATGCCAAATTGCTAGAGCTTATGGCTTAAAGATTTTGGGCACTGCTGGTACTGAGGAAGGA  565

 Q  K  I  V  L  Q  N  G  A  H  E  V  F  N  H  R  E  V  N  Y  I  D  K  I  K  K  Y   211
CAAAAGATTGTTTTGCAAAATGGAGCCCATGAAGTGTTCAATCACAGAGAAGTGAATTACATTGATAAAATTAAGAAGTAT  646

 V  G  E  K  G  I  D  I  I  I  E  M  L  A  N  V  N  L  S  K  D  L  S  L  L  S  H   238
GTTGGTGAGAAAGGAATTGATATAATTATTGAAATGTTAGCTAATGTAAATCTTAGTAAAGACTTGAGTCTTCTGTCACAT  727

 G  G  R  V  I  V  V  G  S  R  G  T  I  E  I  N  P  R  D  T  M  A  K  E  S  S  I   265
GGAGGACGAGTGATAGTTGTTGGCAGCAGAGGTACTATTGAAATAAACCCACGAGACACCATGGCAAAGGAGTCGAGTATA  808

 I  G  V  T  L  F  S  S  T  K  E  E  F  Q  Q  Y  A  A  A  L  Q  A  G  M  E  I  G   292
ATTGGAGTTACTCTCTTTTCCTCAACCAAGGAGGAATTTCAGCAATATGCAGCAGCCCTTCAAGCTGGAATGGAAATTGGC  889

 W  L  K  P  V  I  G  S  Q  Y  P  L  E  K  V  A  E  A  H  E  N  I  I  H  G  S  G   319
TGGTTGAAACCTGTGATAGGTTCTCAATATCCATTGGAGAAGGTGGCCGAGGCTCATGAAAATATCATTCATGGTAGTGGG  970

 A  T  G  K  M  I  L  L  L  *                                                      328
GCTACTGGAAAAATGATTCTTCTCTTATGATGATTAATTCTTTCATGGATTTCCTATGTAATTAGAGGTACTGTCTTTCCC 1051

CCAGTTGTACTTACCCTATCTTTTCTTTAATTAACATTCGATTCCATGAGCTTCTTATGTGAAAAAATAAGATTTTTCTTT 1132

AGAGAGCAGAAGCAGAAGAGTAAAATTTATTGTATAGCTAGCAATATTTTTTTATGCCATCTGTCTCAAATCAAAGAGTCA 1213

TCATAGTAGGAAATAACATGTTAGTTGTCATTTGGCATGAGTGTGCATTCCAGTAATTCTTAATTGATATTTGATTAATTC 1294

CATACCTTTGATTAAAACATGCTAGTTCAAAATAAGACTGCTCAGTTTCCAAGGGTTTTCAAGCCTACTTACCTTTATAAA 1375

GGTTCTCTAGTCTCTGATTAGCCATGACTGTATTGGACTTTGAACATTTTCTGAACTAAAAACCTCTATTCTAAACTAATC 1456

TCATTTGGATGTGTAAGTCTTTTGTAAAGGCAAGAATAAATAATATCCAGGACAATTTATTAGTTTTCTCAGTATTTTCCC 1537

AAATATTAGAATATTTACTTCATTATTGGTTGGCTGCCAATGACCCCATATGTTCTGTGAGAATAGTAGCTTTATCTTTGA 1618

TATAATACATAGTCTCCAAATAGGTAATACTTCGCAATTGATTAGATTTTCAGAGTAGATTTAGAGTTATCTGTTTTTCTG 1699

GTGAGGGTCAAATATTTTTGTTAATTAAGCTCACAAATTTGATAAATTAAGAATTATCTGCATTTGTCTGGTAACATAATA 1780

ATGTGTAATAAAGTCT                                                                   1796
```

FIG. 1. **Nucleotide and deduced amino acid sequences of ζ-crystallin clone HL103.** The single-letter symbol for each amino acid is placed above the central nucleotide of its codon. The termination codon (TGA) is labeled with an asterisk (*). The two putative polyadenylation signals are underlined.

998-1000 followed by a second TGA triplet. Although no poly(A) tail was present in the clone, two putative polyadenylation signals (AATAAA) were found at nucleotides 1491-1496 and 1787-1792. The clone extended only four nucleotides downstream of the latter which is located in the same relative position as a functional polyadenylation signal in the guinea pig messenger RNA from lens (6) (Figs. 1 and 2).

The open reading frame was 984 nucleotides in length, and coded for a protein with the same number of residues as that in guinea pig (328 amino acids) and with a calculated molecular mass and isoelectric point of 34910 daltons and 8.73, respectively. The total number of residues differing from the guinea pig sequence was 52, which corresponds to 84% identity between the two species (Fig. 3).

The analysis of the deduced amino acid sequence of ζ-crystallin in guinea pig established a phylogenetic and structural relationship between this protein and the family of zinc-containing alcohol/polyol dehydrogenases (ADH) (6). The residues associated with the maintenance of the tertiary structure of the ADH molecule were strongly conserved in ζ-crystallin from guinea pig. Of the 13 residues strictly conserved in the ADH family as well as in guinea pig ζ-crystallin, only the Val in position 83 is substituted by Ala in the human sequence. Similarly, there is only one substitution (Thr>Ser in position 266) among the 26 amino acids conserved in the majority of the ADH sequences compared. This result indicates an apparent conservation of the structure of ζ-crystallin in humans and in guinea pig. As in guinea pig, the residues binding the catalytic zinc atom in alcohol dehydrogenases (6) are absent in the human sequence consistent with the enzymatic data (4) showing that ζ-crystallin does not have alcohol dehydrogenase activity.

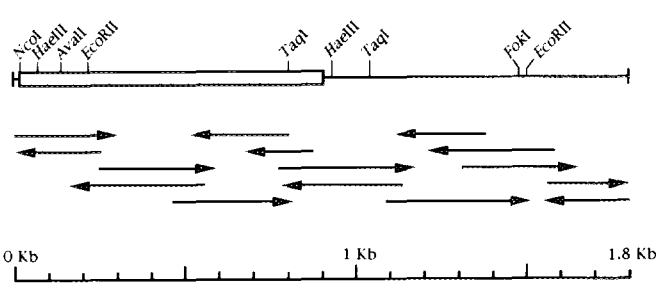

FIG. 2. **Restriction map and sequencing strategy used in the analysis of the cDNA clone encoding ζ-crystallin.** The coding region is represented by an open box. Noncoding regions are solid lines. Arrows show the direction and extent to which sequence was determined. The restriction map reveals only relevant or unique sites.

```
GP   ATGQKLMRAI RVFEFGGPEV LKVQSDVAVP IPKDHQVLIK VHACGINPVE
     ::::::::::. :::::::::: ::..::.::: :::::::::: :::::.::::
H    ATGQKLMRAV RVFEFGGPEV LKLRSDIAVP IPKDHQVLIK VHACGVNPVE
              10          20          30          40          50

GP   TYIRSGTYTR IPLLPYTPGT DVAGVVESIG NDVSAFKKGD RVFTTSTISG
     :::::::::.: :::::::::. :::::.:...: .. :::::::: ::::.::::: 
H    TYIRSGTYSR KPLLPYTPGS DVAGVIEAVG DNASAFKKGD RVFTSSTISG
              60          70          80          90          100

GP   GYAEYALASD HTVYRLPEKL DFRQGAAIGI PYFTACRALF HSARAKAGES
     :::::::::.: :::::.::::: ::.::::::: :::::.::: ::: .:::::
H    GYAEYALAAD HTVYKLPEKL DFKQGAAIGI PYFTAYRALI HSACVKAGES
              110         120         130         140         150

GP   VLVHGASGGV GLAACQIARA YGLKVLGTAG TEEGQKVVLQ NGAHEVFNHR
     :::::::::: :::::::::: ::::.::::: :::::::::: ::::::::::
H    VLVHGASGGV GLAACQIARA YGLKILGTAG TEEGQKIVLQ NGAHEVFNHR
              160         170         180         190         200

GP   DAHYIDEIKK SIGEKGVDVI IEMLANVNLS NDLKLLSCGG RVIIVGCRGS
     ...:::.::: .::::.:.: :::::::::: .::.::: :: :::.::.::: 
H    EVNYIDKIKK YVGEKGIDII IEMLANVNLS KDLSLLSHGG RVIVVGSRGT
              210         220         230         240         250

GP   IEINPRDTMA KESTISGVSL FSSTKEEFQQ FASTIQAGME LGWVKPVIGS
     :::::::::: :::.: ::.: :::::::::: .:...::::: .::.::::::
H    IEINPRDTMA KESSIIGVTL FSSTKEEFQQ YAAALQAGME IGWLKPVIGS
              260         270         280         290         300

GP   QYPLEKASQA HENIIHSSGT VGKTVLLM
     ::::::...: ::::::.::. .:: .::.
H    QYPLEKVAEA HENIIHGSGA TGKMILLL
              310         320      328
```

FIG. 3. **Alignment of human (H) and guinea pig (GP) (5) ζ-crystallin amino acid sequences.** The identical residues are connected by double dots (:) and conservative substitutions are indicated by a single dot (.). The 33 amino acids of the coenzyme-binding segment extensively conserved among dehydrogenases (12) are underlined.

Concordant with the importance of NADPH binding to the ability of ζ-crystallin to carry out its role as a quinone reductase (4), the region that constitutes the backbone of the coenzyme-binding domain of the dehydrogenases (11) shows a clear degree of conservation. Only one amino acid in this 33 residue segment differs in the human sequence relative to guinea pig (the amino acid Ile at residue 175 instead of Val).

Although its physiological function remains unclear, the presence of ζ-crystallin at low levels in lenses from various species implies that it does have a catalytic function in lens aside from its presumed structural role in those species where it is expressed as a crystallin. The results reported here are consistent with this idea and provide the first step toward analysis of the structure of the human ζ-crystallin gene. Comparison of the promoters and other elements regulating the expression of this gene in species where it is present as a crystallin with species where it is expressed only at low levels in the lens should help clarify the evolutionary process whereby this protein was recruited as a lens crystallin.

REFERENCES

1. Huang, Q.-L., Russell, P., Stone, S.H., and Zigler, J.S., Jr. (1987) Curr. Eye Res. 6, 725-732.

2. Garland, D., Rao, P.V., Corso, A.D., Mura, U., and Zigler, J.S., Jr. (1991) Arch. Biochem. Biophys. 285, 134-136.

3. Rao, P.V., and Zigler, J.S., Jr. (1992) Biochim. Biophys. Acta 1117, 315-320.

4. Rao, P.V., Krishna, M.C., and Zigler, J.S., Jr. (1992) J. Biol. Chem. 267, 96-102.

5. Rodokanaki, A., Holmes, R.K., and Borrás, T. (1989) Gene 78, 215-224.

6. Borrás, T., Persson, B. and Jörnvall, H. (1989) Biochemistry 28, 6133-6139.

7. Rodriguez, I.R., Gonzalez, P., Zigler, J.S., Jr. and Borrás, T. (1992) Biochim. Biophys. Acta 1180, 44-52.

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

9. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. (U.S.A) 74, 5463-5467.

10. Kozak, M. (1991) J. Biol. Chem. 266, 19867-19870.

11. Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.O., Tapia, O., Brändén, C.I., and Åkerson, Å. (1976) J. Mol. Biol. 102, 27-59.