Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 6876–6880, November 1982
Biochemistry

Extensive intragenic sequence homology in two distinct rat lens γ-crystallin cDNAs suggests duplications of a primordial gene

(cDNA cloning/DNA sequence determination/gene generation)

ROB J. M. MOORMANN*, JOHAN T. DEN DUNNEN*, HANS BLOEMENDAL†, AND JOHN G. G. SCHOENMAKERS*

Departments of *Molecular Biology and †Biochemistry, University of Nijmegen, Nijmegen, 6525 ED, The Netherlands

Communicated by L. N. M. Duysens, August 17, 1982

ABSTRACT **The nucleotide sequences of two different rat lens γ-crystallin cDNA clones, pRLγ2 and pRLγ3, have been determined. pRLγ3 contains the complete coding information for a γ-crystallin of 173 amino acids whereas pRLγ2 is incomplete in that it lacks the codons for the first three amino acids of a separate but very homologous γ-crystallin of identical length. Both rat γ-crystallins are homologous to the known amino acid sequence of bovine γ-crystallin II which is only a single amino acid longer. The length of the region downstream the coding sequence to the A-A-T-A-A-A polyadenylylation signal sequence is 40 nucleotides in each clone. In pRLγ3 the poly(A) signal sequence is followed at 14 nucleotides by a remnant of the poly(A) tail which indicates that this clone contains a complete 3′ noncoding region. pRLγ2 has only seven nucleotides following this signal sequence and no poly(A) tail, suggesting an incomplete 3′ end. The cDNA clones show an overall nucleotide sequence homology of 85%. The mutual homology at the amino acid level is 73% whereas their amino acid homology with bovine γ-crystallin II is about 70%. The nucleotide sequence of each clone also reveals a high intragenic homology and seems to be duplicated in itself. We suggest that the γ-crystallin genes have arisen by multiple duplications of a primordial gene which consisted of about 120 nucleotides.**

The transparency and high refractive index of the normal eye lens necessary for focusing visible light on the retina is achieved not only by a regular arrangement of the lens fiber cells during growth of the lenticular body but also by the extremely high concentration and the supramolecular organization of the lens-specific proteins, the crystallins, within each fiber cell. In the mammalian lens, three major classes of crystallins are distinguished: α, β, and γ. The largest oligomeric protein, α-crystallin, is composed of two primary gene products αA_2 and αB_2 of M_r 20,000 which, together with their deamidation products αA_1 and αB_1, associate to aggregates with an average M_r of 800,000 (1–3). The β-crystallins are an even more heterogeneous class of oligomeric proteins; at least five distinct polypeptides are involved in their association and the M_rs range from 20,000 to 32,000 (4, 5). The γ-crystallins, which contribute about 40% of the total protein mass in rat and mouse lens (6, 7), are monomeric and have a M_r of ≈20,000 (8–10). At least five closely related γ-crystallins have been identified in the bovine (8) and rat lens (11).

The γ-crystallins are rich in cysteine compared with the other water-soluble lens proteins α- and β-crystallins (10, 12). The most prominent species in the bovine lens, γ-crystallin II, has six free sulfhydryl groups (9). Although it was shown in the rodent lens that, with aging, there is an almost complete oxidation (SH to S—S) of lens protein (13, 14), little is known of the conformation and intermolecular interactions of the crystallins in the lens fiber cells *in situ*. However, it has been established that increased light scattering of senile nuclear cataract is associated with denaturation, chemical crosslinking including disulfide bridges, and aggregation of crystallins (15).

Recently, the three-dimensional structure of a lens protein, bovine γ-crystallin II, has been determined by Blundell *et al.* (16). It turned out that the sulfhydryl groups in the highly symmetric two-domain β structure of this protein were found in positions that would allow both intra- and intermolecular crosslinking. Disulfide bridges can be formed through some sulfhydryls without destabilization or denaturation of the protein, but denaturation would certainly result if the sulfhydryls that are buried in hydrophobic regions were covalently crosslinked. Moreover, a large decrease in the γ-crystallin content of the lens has been demonstrated in hereditary cataracts (see ref. 7 and citations therein).

These findings and the observations that the γ-crystallins are expressed exclusively in the cortical fibers (17) in which they are synthesized at a remarkably decreased rate upon aging (18, 19) prompted us to a more detailed analysis of the γ-crystallin genes as a first step toward a better understanding of their expression in normal and disordered lens. Few studies have dealt with crystallin genes (20–22). We previously described the cloning and characterization of double-stranded DNA copies of various α-, β-, and γ-crystallin-specific rat lens mRNAs (11) and determined the nucleotide sequence of most of the mRNA encoding the αA_2-crystallin chain (23). In this study we describe the nucleotide sequence of two mRNAs coding for two highly homologous rat γ-crystallins.

MATERIALS AND METHODS

Enzymes and Chemicals. All restriction endonucleases used were purchased from either Boehringer Mannheim or Bethesda Research Laboratories. Polynucleotide kinase was from P-L Biochemicals. Calf intestine alkaline phosphatase and *Escherichia coli* DNA polymerase I (large fragment) were from Boehringer Mannheim. [γ-^{32}P]ATP (3,000 Ci/mmol; 1 Ci = 3.7 × 10^{10} becquerels) and [α-^{32}P]dCTP (3,000 Ci/mmol) were from Amersham. The synthetic oligonucleotide primer 5′ d(G-T-A-A-A-A-C-G-A-C-G-G-C-C-A-G-T-G)3′ which is complementary to 18 bases directly preceding the region containing multiple restriction sites in M13mp7 (24) was constructed by J. van Boom and co-workers (Department of Organic Chemistry, Leiden University).

DNA Sequence Analysis. The isolation of recombinant DNA and the determination of nucleotide sequences by the method of Maxam and Gilbert (25) were carried out essentially as described (23). In some cases, ambiguities in the nucleotide sequence of pRLγ3 could not clearly be solved by this way of sequence determination. Therefore, *Pst* I fragments of this clone

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.

were inserted in *Pst* I-linearized, alkaline phosphatase-treated M13mp7 (24) and transfected in *E. coli* JM103 (24). Single-stranded DNA from colorless plaques was subsequently analyzed by the chain-termination method of DNA sequence determination as described by Sanger *et al.* (26) with the 18-base oligonucleotide mentioned above as primer.

RESULTS

After restriction enzyme analysis and positive hybridization selection, the cDNA clones pRLγ2 and pRLγ3 and another one designated pRLγ1 were selected as appropriate candidates for full-length transcripts (650–750 base pairs) of the 10S mRNA class in which the γ-crystallin mRNAs are dominantly present (11). Refinement of the physical maps and determination of the nucleotide sequences indicated that pRLγ1 and pRLγ2 were identical except for the 5′ and 3′ termini which were slightly shorter in pRLγ1 (data not shown).

Fig. 1 shows the physical maps of pRLγ2 and pRLγ3. Although these clones had several restriction cleavage sites in common at identical positions (only those used in determining the DNA sequence are indicated), pRLγ3 was distinctly different from pRLγ2 in that the former had one extra *Pst* I site and two extra *Hin*fI sites.

It should be noted that a positive hybridization assay cannot discriminate between these clones because the same set of γ-crystallins as translation products show up with each clone (11). From this we inferred that pRLγ2 and pRLγ3 are different but highly homologous γ-crystallin-specific cDNA clones. This assumption is clearly sustained by the nucleotide sequences of the cDNAs (Fig. 2). At this level the clones show an overall homology of 85%. However, the homology varies in subregions; up to nucleotide 250 it is 90%, from nucleotide 250 to 340 it diminishes to 60%, and from nucleotide 340 to 580 it increases again to 80%.

The nucleotide sequences of pRLγ2 and pRLγ3 contain open reading frames encoding polypeptides of 170 and 173 amino acids, respectively. The amino acid sequences indicate

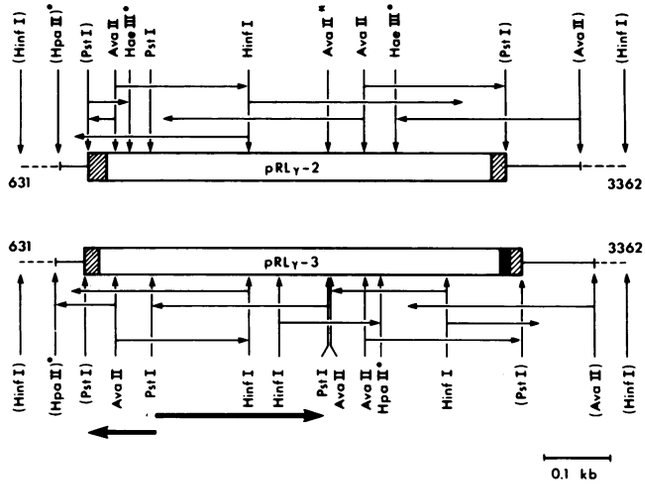

FIG. 1. Restriction map of pRLγ2 and pRLγ3, showing the restriction sites used for sequence analysis. The thin horizontal arrows indicate the orientation and approximate length of sequences obtained by the chemical degradation method (25); the thick arrows indicate orientation and length of sequences deduced by the dideoxynucleotide method (26). The cleavage coordinates of the *Hin*fI sites flanking the *Pst* I site of pBR322 are indicated. Open boxes, γ-crystallin mRNA sequences; hatched boxes, poly(dG-dC) tracks; black box in pRLγ3, remnant of the poly(A) track of the mRNA template used in cDNA synthesis; asterisk, a methylated *Ava* II site.

that, compared with pRLγ3, pRLγ2 lacks three amino acid residues at its NH_2 terminus. pRLγ3 most probably comprises the complete coding information contained in the γ-crystallin mRNA. This is deduced from comparison with the known amino acid sequence of γ-crystallin II from bovine origin (9, 16). The sequence of this protein starts with a glycine and is completely homologous with the first 13 amino acids of pRLγ3 (Fig. 2). In contrast to earlier findings by Zigman *et al.* (27) who reported alanine as the NH_2 terminus, Wagner and Fu (12) showed by Edman degradation analysis that rat crystallins only have glycine at their NH_2 terminus. In line with this is our finding of T-G preceding the NH_2-terminal glycine codon in pRLγ3. From extended analysis of genomic clones, which will be presented elsewhere, it turned out that in pRLγ3 the A of the initiator ATG codon is missing.

Comparison of the two rat clones at the amino acid level indicates that 60% of the changes in amino acid sequences were due to two or more nucleotide changes in a single codon. More or less nonconservative changes are apparent at amino acid position 15 (Ser-His), 22 (Cys-His), 23 (Pro-Ser), 27 (Thr-Pro), 53 (His-Cys), 100 (Val-Gln), 109 (Ser-Pro), 110 (Cys-His), 148 (Gln-Gly), and 168 (Ser-Arg). The mutual homology at the amino acid level between the two clones is 73%. The amino acid homology with bovine γ-crystallin II is 66% for pRLγ2 and 70% for pRLγ3. The bovine γ-crystallin II has one extra amino acid compared to the two rat γ-crystallins.

Another interesting feature is the high content of cysteine residues in the γ-crystallins. Although there is no evidence for disulfide bridges in the mature protein (10), it is known that the location of these residues is often highly conserved in protein. If the NH_2-terminal glycine of pRLγ3 is taken as the first amino acid, Cys residues 18, 32, and 108 are at the same positions in pRLγ2, pRLγ3, and bovine crystallin II as well and Cys-41 is the same only in the two rat clones. pRLγ2 has additional nonshared cysteine residues at positions 22, 110, and 129 whereas pRLγ3 has an extra cysteine at position 53. Bovine crystallin II has two additional cysteine residues at positions 15 and 23. It has been shown (13, 14) that, during the process of aging of the rat lens, almost all soluble protein SH is oxidized to S—S. Because the γ-crystallins are rich in cysteines compared to α- and β-crystallin (10, 12, 28), it is mainly these proteins that should undergo inter- and intramolecular crosslinking during this process.

In the 3′ noncoding regions in the rat cDNA clones we find a stretch of 14 A residues at 22 nucleotides after the A-A-T-A-A-A polyadenylylation signal sequence (29) in pRLγ3. The poly(A) stretch being a remnant of the oligonucleotide primer used in the synthesis of the cDNA indicates that this clone contains the complete 3′ end of the mRNA. No poly(A) stretch was found in pRLγ2. In this clone the poly(dC) tail used in cloning started 11 nucleotides after the polyadenylylation signal sequence. In view of this and the high degree of homology between pRLγ2 and pRLγ3, we anticipate that approximately 10–15 nucleotides of the mRNA are missing at the 3′ end of this clone. The homology between the short (65 nucleotides for pRLγ3) 3′ noncoding regions is 75%, which is only slightly lower than the overall homology between these clones. This feature is in agreement with the retained homology in the short 3′ noncoding regions of 52 and 53 nucleotides, respectively, of two rat preproinsulin genes (30) but contrasts with findings in other families of closely related genes like the human leukocyte interferons (31) and several human (32) and mouse β-globin-like genes (33) in which the 3′ noncoding regions diverge more rapidly than the protein coding regions.

Experiments involving *in vitro* translation of mRNAs with large 3′ deletions suggest that intact 3′ noncoding regions are

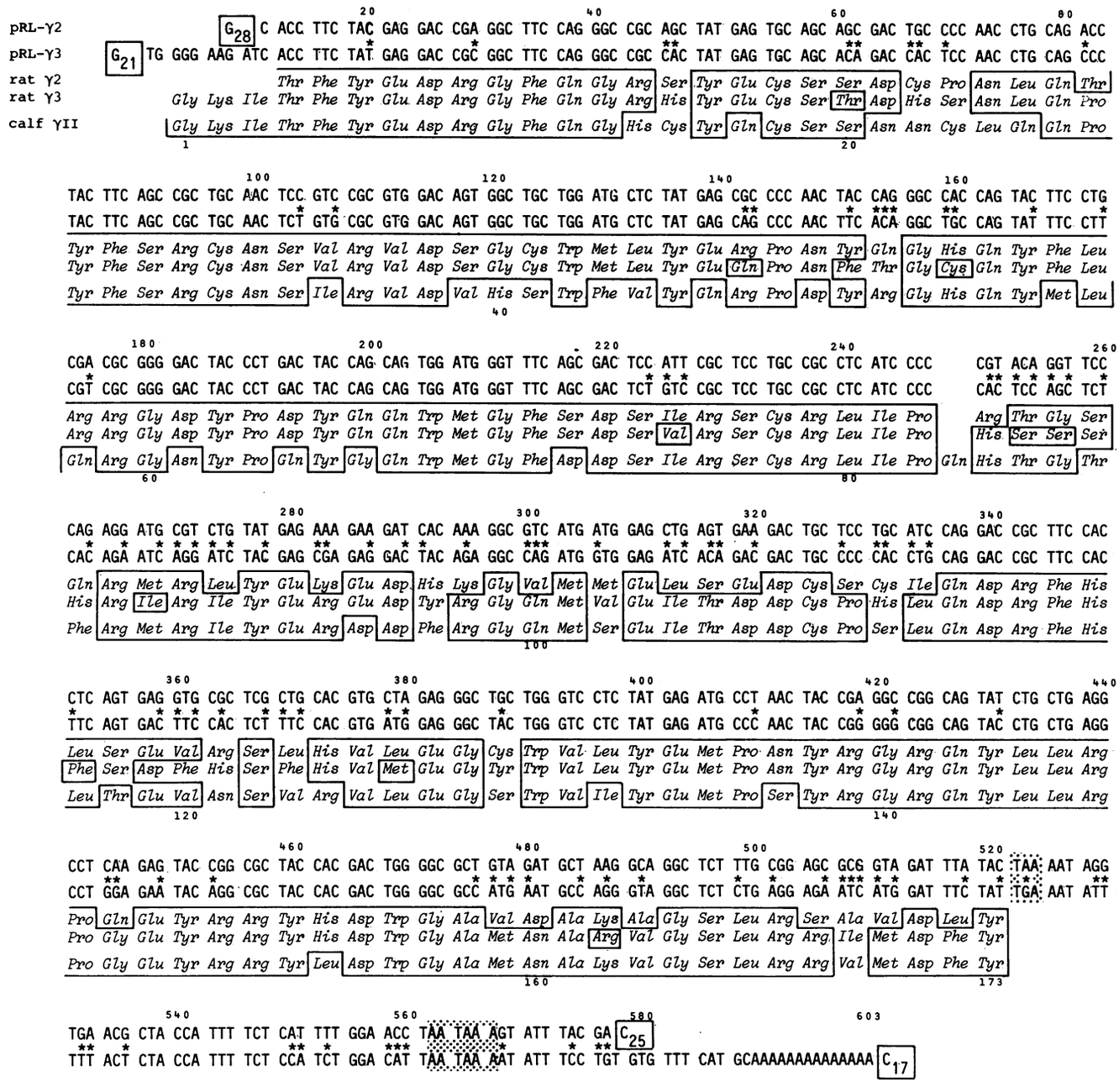

FIG. 2. Nucleotide sequence of the rat lens γ-crystallin cDNA inserts in pRLγ2 and pRLγ3 and the deduced amino acid sequence. For comparison, the amino acid sequence of bovine γ-crystallin II (9) as modified by Blundell *et al.* (16) is also given. Stippled area, nucleotide sequence involved in poly(A) addition; asterisks, differences in the nucleotide sequence between the rat γ-crystallin cDNAs; boxes, regions with identical amino acid sequences in both rat γ-crystallins and bovine γ-crystallin II.

not needed for efficient translation of the protein encoding regions in the mRNA (34, 35). Furthermore, we found that the αA_2-crystallin mRNA has a 3′ noncoding region of 520 nucleotides which almost equals the length of the coding region of this mRNA (23). However, no specific function could be assigned to this large noncoding region. These findings suggest that the length and the precise sequence of much of the 3′ noncoding regions of these eukaryotic mRNAs are not essential for function. It remains to be seen, however, whether the 3′ noncoding regions in the other not yet analyzed γ-crystallin genes are as conserved as those in pRLγ2 and pRLγ3. If so, this might reflect a functional importance of these sequences in the γ-crystallin multigene family.

DISCUSSION

Close analysis of the two γ-crystallin cDNA sequences shows that they have an axis of symmetry which divides them into NH_2- and COOH-terminal halves. A two-by-two arrangement of these halves (Fig. 3) yields regions *a*–*a*′ and *b*–*b*′; the latter reveals the higher homology. In order to get a higher degree of homology between the *a* and *a*′ regions of pRLγ2 and pRLγ3, a six-nucleotide deletion was introduced following the C at position 83 in the *a* region of both cDNA sequences. For the same reason, a trinucleotide was deleted at position 311 in the *a*′ regions of pRLγ2 and pRLγ3. However, this needs to be compensated for six nucleotides upstream in the corresponding *a* region by deletion of a trinucleotide behind the nucleotide

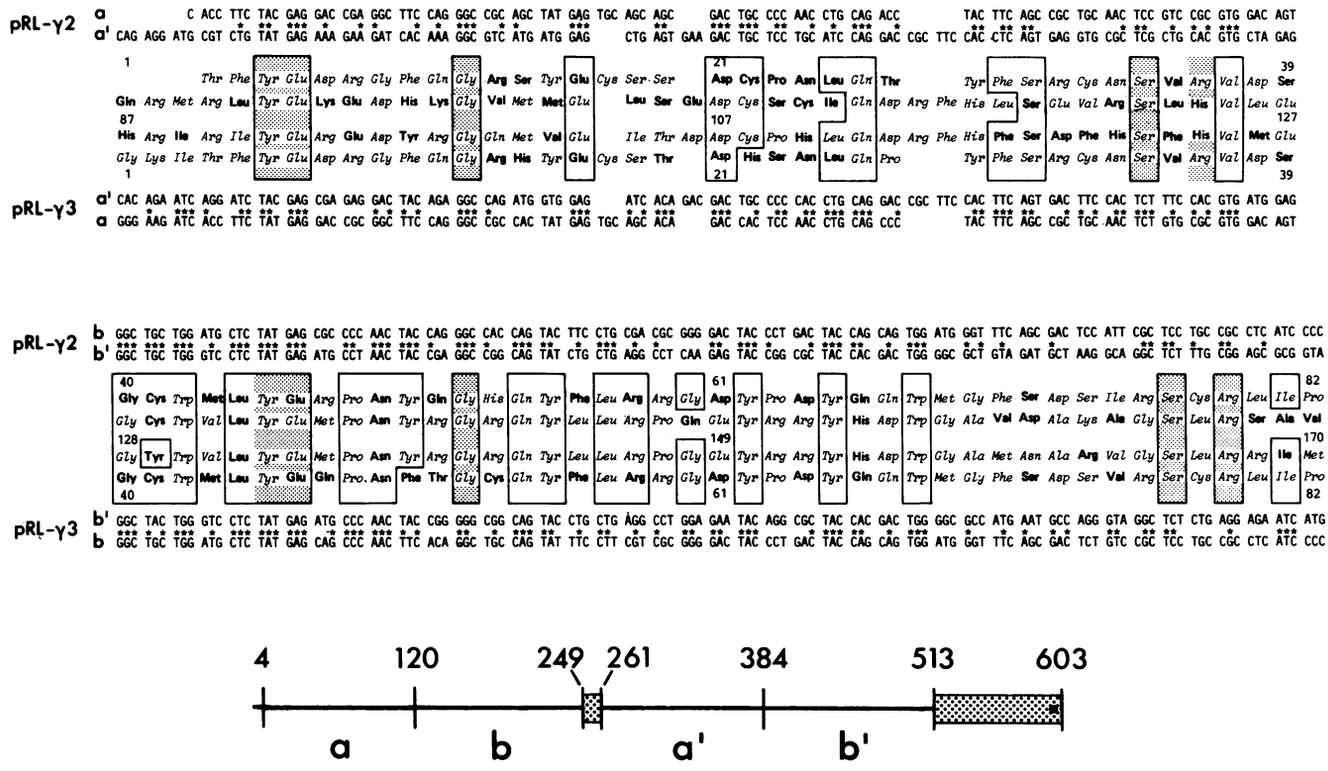

FIG. 3. Intragenic nucleotide sequence homology in the rat lens γ-crystallin genes. The nucleotide sequence of pRLγ2 and pRLγ3 has been divided into subregions (*a*, *a'*, *b*, and *b'*) as indicated below the sequence. Arabic numerals represent the nucleotide positions as given in Fig. 2. Region 249–261 represents the "connecting peptide" (see text). The subregions and their corresponding amino acid sequences have been arranged in such a way that a mutual comparison can be made between *a* and *a'* (or *b* and *b'*) of either pRLγ2 or pRLγ3 and also between these subregions in both γ-crystallin sequences. Nucleotide sequence homology in each subregion is indicated by asterisks. Amino acids in each rat γ-crystallin sequence that differ from the bovine γ-crystallin II amino acid sequence (cf. Fig. 2) are given in boldface. Boxed areas, amino acid residues that are identical in *a* and *a'* (or *b* and *b'*) of both rat γ-crystallins; stippled areas, amino acid residues located at topologically equivalent positions in all subregions of rat γ-crystallin and bovine γ-crystallin II as well.

at position 62. It is possible to increase homology by one base in this small region of pRLγ3 by placing the deletion in the *a* stretch six nucleotides upstream from where it is placed now. We prefer the former sequence alignment because the other way of matching would introduce a deleted amino acid in the *a* strain of pRLγ3 which would be two positions further than the deletion in pRLγ2. This, in turn, should imply that the trinucleotide deletions at these positions in the two genes would have occurred after duplication of the ancestral γ-crystallin gene. We favor, however, the idea that the deletions occurred as indicated and before duplication of the ancestral gene.

The overall degree of homology, as deduced from Fig. 3, between regions *a* and *a'* is 44% for pRLγ2 and 52% for pRLγ3. This is lower than the overall homology between regions *b* and *b'*, 55% for pRLγ2 and 56% for pRLγ3. Because the deletions introduced in regions *a* and *a'* only enhance the mutual homology between these regions, the actual homology difference between *a*/*a'* and *b*/*b'* is even higher and mainly due to the extremely high homology between the first 13 codons in regions *b* and *b'*.

The four residues tyrosine, glutamate, glycine, and serine are conserved at equivalent positions in all the subregions of the two cDNA clones (shaded boxed areas in Fig. 3). From x-ray analysis at 2.6-Å resolution, Blundell *et al.* (16) deduced that bovine γ-crystallin II shows the highest degree of intrachain symmetry of any protein whose three-dimensional structure has been elucidated so far. The protein has a two-domain β-structure folded into four remarkably similar Greek key motifs. These motifs correspond exactly with the subregions *a*, *b*, *a'*, and *b'* as indicated in Fig. 3. The two domains are formed by *a* + *b* and *a'* + *b'* and are connected at *b* and *a'* by a 12-nucleotide region, nucleotide positions 249–260, which codes for the so-called connecting peptide.

The backbone of the intramotif structure in the protein is mainly formed by the four residues mentioned above. The chain folds between Glu-7 and Gly-13. This allows Tyr-6 to pack against Phe-11 which is only nonconservatively varied to His in subregion *a'* of pRLγ2. The other conserved residue is Ser-34 which stabilizes the structure by hydrogen bonding to Phe-11 (or to a topologically equivalent residue). The fifth residue which participates in this unusual foldover feature is Arg-36 which, however, is rather conservatively varied to His in subregion *a'* of pRLγ2 and pRLγ3. The deletion of six nucleotides introduced in subregion *a* of pRLγ2 and pRLγ3 corresponds with hypothetical amino acids located at a position in the protein where Blundell *et al.* (16) deleted three amino acids in bovine γ-crystallin II. The amino acid deletions at position 18 and 21 are located just before and at the end of the second turn in motif I and motif III. It is conceivable that two deletions have been introduced here because of the fact that deletion of only a single amino acid residue would have had such an impact on the structure in this region that a second nearby deletion was necessary. Furthermore, motif II and motif IV (subregions *b* and *b'*) in both rat γ-crystallins are four residues larger than motif I (subregion *a*) which is exactly the same as in bovine γ-crystallin II. The contacts between the two domains are largely hydrophobic. The residues involved in packing the two domains together are methionine and valine at positions 43 and 131, re-

spectively, and phenylalanine and leucine at positions 56 and 144. It is interesting to note that methionine and phenylalanine are in the reversed orientation in bovine γ-crystallin II (cf. Fig. 2).

The connecting peptide which forms the covalent linkage between the two domains consists of four residues (positions 83–86) which are conservatively varied among rat and calf. However, there is a clear difference when the residues directly flanking the connecting peptide are considered. The glutamine preceding the connecting peptide in bovine γ-crystallin II has been deleted in both rat γ-crystallins whereas the phenylalanine following the connecting peptide is rather nonconservatively varied to histidine or glutamine in pRLγ3 and pRLγ2, respectively. How far this feature interferes with interdomain linkage is hard to say. Because these amino acids are next to the connecting peptide which is at the surface of the molecule, it is feasible that there should be some space for amino acid variation without dramatic effects on tertiary structure. It is also conceivable that the reversion of phenylalanine and methionine has taken place to accommodate small three-dimensional structural alterations that could have taken place.

From the results presented in this study it is clear that the γ-crystallins are well conserved among calf and rat at the structural level. Moreover, the structure of the cDNA sequences suggests that the primordial γ-crystallin gene consisted of only one subdomain—say *a*. Duplication of this subregion then generated a domain-like gene which, by tandem duplication, evolved into the mature γ-crystallin gene described above. In this respect, γ-crystallin is not unique and seems to resemble well-characterized genes—for instance, the chicken ovomucoid gene, which has three domains showing 40–60% of homology at the nucleotide level (36), or the chicken collagen gene for which it has been shown that the ancestral gene might have arisen by multiple duplication events of a single genetic unit consisting of only 54 base pairs (37).

The occurrence of different nonallelic γ-crystallin genes was suggested in 1978 (10). Our recent data on restricted genomic DNA and subsequent molecular cloning demonstrate that there are at least five γ-crystallin genes in rat. This indicates that, besides the intragenic type of duplication, single or tandem duplications of mature γ-crystallin sequences also have taken place more recently to generate a multigene family.

A considerable increase in the resolution to 1.9 Å has meanwhile been achieved in x-ray studies (T. Blundell and C. Slingsby, personal communication). This enabled them to refine further the previously reported (9, 16) calf γII-crystallin amino acid sequence. Therefore, the degree of homology between the calf and rat γ-crystallins as presented in this paper must be considered as a minimum.

This work was supported by The Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

1. Van der Ouderaa, F. J., de Jong, W. W. & Bloemendal, H. (1973) *Eur. J. Biochem.* **39,** 207–222.
2. Van der Ouderaa, F. J., de Jong, W. W., Hilderink, A. & Bloemendal, H. (1974) *Eur. J. Biochem.* **49,** 157–168.
3. Spector, A. & Katz, E. (1965) *J. Biol. Chem.* **240,** 1979–1985.
4. Van Kleef, F. S. M., de Jong, W. W. & Hoenders, H. J. (1975) *Nature (London)* **258,** 264–266.
5. Siezen, R. J., Bindels, J. G. & Hoenders, H. J. (1978) *Eur. J. Biochem.* **91,** 387–396.
6. Ochen, P. R., Fu, S. C. J., Hart, R., White, J. H., Wagner, B. J. & Lewis, K. F. (1977) *Exp. Eye Res.* **24,** 355–367.
7. Wada, E., Sugiura, T., Nakamura, H. & Tsumita, T. (1981) *Biochim. Biophys. Acta* **667,** 251–259.
8. Björk, I. (1970) *Exp. Eye Res.* **9,** 152–157.
9. Croft, L. R. (1972) *Biochem. J.* **128,** 961–970.
10. Slingsby, C. & Croft, L. R. (1978) *Exp. Eye Res.* **26,** 291–304.
11. Dodemont, H. J., Andreoli, P. M., Moormann, R. J. M., Ramaekers, F. C. S., Schoenmakers, J. G. G. & Bloemendal, H. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5320–5324.
12. Wagner, B. J. & Fu, S. C. J. (1978) *Exp. Eye Res.* **26,** 255–265.
13. East, E. J., Chang, R. C., Yu, N. T. & Kuck, J. F. R. (1978) *J. Biol. Chem.* **253,** 1436–1441.
14. Askren, C. C., Yu, N. T. & Kuck, J. F. R. (1979) *Exp. Eye Res.* **29,** 647–654.
15. Harding, J. (1981) in *Molecular and Cellular Biology of the Eye Lens,* ed. Bloemendal, H. (Wiley, New York), pp. 327–365.
16. Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. & Wistow, G. (1981) *Nature (London)* **289,** 771–777.
17. Papaconstantinou, J. (1965) *Biochim. Biophys. Acta* **107,** 81–90.
18. Slingsby, C. & Croft, L. R. (1973) *Exp. Eye Res.* **17,** 369–376.
19. Bours, J., Doepfmer, K. & Hockwin, O. (1976) *Doc. Ophthalmol. Proc. Ser.* **8,** 75.
20. Bhat, S. P., Jones, R. E., Sullivan, M. A. & Piatigorsky, J. (1980) *Nature (London)* **284,** 234–238.
21. Tomarev, M., Dolgilevich, S. M., Kozlov, K. A., Zinovieva, R. D., Dzhumagaliev, E. B., Kogan, G. L., Skobelev, N. A., Mikhailov, A. T., Frovola, L. & Gause, G. G. (1982) *Gene* **17,** 131–138.
22. King, C. R., Shinohara, T. & Piatigorsky, J. (1982) *Science* **215,** 985–987.
23. 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.
24. Messing, J. M., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9,** 309–321.
25. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
27. Zigman, S., Schultz, J. & Yulo, T. (1970) *Biochim. Biophys. Acta* **221,** 576–582.
28. Croft, L. R. (1980) *Proc. Int. Soc. Eye Res.* **1,** 118 (abstr.).
29. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
30. Lomecido, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. & Tizard, R. (1979) *Cell* **18,** 545–558.
31. Goeddel, D. V., Leung, D. W., Dull, T. J., Gross, M., Lawn, R. M., McCandliss, R., Seeburg, P. H., Ulrich, A., Yelverton, E. & Gray, P. W. (1981) *Nature (London)* **290,** 20–26.
32. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, A. R., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. R., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21,** 653–668.
33. Konkel, D. A., Maizel, J. V. & Leder, P. (1979) *Cell* **18,** 865–873.
34. Kronenberg, H. M., Roberts, B. E. & Efstratiadis, A. (1979) *Nucleic Acids Res.* **6,** 153–166.
35. Nunberg, J. H., Kaufman, R. J., Chang, A. C. Y., Cohen, S. N. & Schminke, R. T. (1980) *Cell* **19,** 355–364.
36. Stein, J. P., Catterall, J. F., Kristo, P., Means, A. R. & O'Malley, B. W. (1980) *Cell* **21,** 681–687.
37. Yamada, Y., Anvedimento, V. E., Mudryj, M., Ohkulo, H., Vogeli, G., Irani, M., Pastan, I. & De Crombrugghe, B. (1980) *Cell* **22,** 887–892.