BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **240,** 51–56 (1997)
ARTICLE NO. RC977600

Characterization of γS-Crystallin Isoforms from Lip Shark (*Chiloscyllium colax*): Evolutionary Comparison between γS and β/γ Crystallins[1]

Fu-Ming Pan,* Ming-Hong Chuang,† and Shyh-Horng Chiou*,†,[2]
*Laboratory of Crystallin Research, †Institute of Biochemical Sciences, National Taiwan University, and *Institute of Biological Chemistry, Academia, P.O. Box 23-106, Taipei, Taiwan*

Received September 30, 1997

γS-Crystallin from shark eye lenses, formerly termed βs crystallin in mammalian lenses, is structurally characterized in this study by cDNA cloning and sequencing. To facilitate sequence characterization of γS-crystallin possessing intermediate structural properties between β- and γ-crystallins, cDNA mixture was constructed from the poly(A)$^+$mRNA isolated from shark eye lenses, and amplification by polymerase chain reaction (PCR) was carried out to obtain nucleotide segments encoding multiple shark γS-crystallins. Sequencing several positive clones revealed that a multiplicity of isoforms exists in the γS-crystallin class of this cartilaginous fish, similar to authentic γ-crystallin family characterized from the same shark species. Comparison of protein sequences encoded by two representative shark γS1 and γS2 cDNAs with those published sequences of β-, γ-, and γS crystallins from bovine, human, bullfrog and carp lenses indicated that there is about 35–64% sequence homology between shark γS crystallins and structurally related crystallins from different evolutionary classes, with a higher sequence similarity between shark γS and mammalian γ-crystallins than that of shark γS and carp γS or bovine γS crystallins. A phylogenetic tree constructed on the basis of the sequence divergence among various β-, γ-, and γS crystallins corroborates the closer relatedness of shark γS to authentic γ-crystallin than to mammalian and teleostean γS crystallins. It further strengthens the supposition that ancestral precursors of γS-crystallins were present in the shark lens long before the appearance of present-day teleostean and mammalian γS-crystallins.

The abundant presence of various common and specific classes of structurally conserved proteins in eye lenses (crystallins) of different species of vertebrates constitutes a good model system to unravel the complex process of evolution in structurally homologous proteins (1,2). Fish represents the oldest and most diverse group of vertebrates. The modern fishes comprise two major classes of piscine species: [1] Osteichthyes or teleostean (bony) fishes and [2] Chondrichthyes or cartilaginous fishes (sharks and skates).

The study of lens crystallins from the lowest piscine class is of special interest from the evolutionary point of view because they constitute the early protein forms of vertebrates and are thought to have been ancestral to those of land vertebrates. Sharks diverged from the Placodermi (one class of armored fishes) long before the appearance of modern bony fishes and amphibians (3). The characterization of shark crystallins would be deemed very important for the phylogenetic comparison in light of the recent elucidation of the complete sequences of γ-crystallins from several species of teleostean fishes in *Osteichthyes* (4–6).

The present study was performed in the endeavor to have a structural characterization of one unique class of lens crystallins with their primary structure lying between the well-known β- and γ-crystallins. This class of crystallin, formerly called βS and now renamed γS crystallin (7,8), exists as a monomeric protein which is similar to the major authentic γ-crystallins. However unlike γ-crystallins which possess a free N-terminal amino-acid residue, γS-crystallin has a blocked N terminus as most members of β-crystallin family. In this report we have for the first time cloned and sequenced γS crystallins from the shark eye lenses, which possess some structural properties distinguishable from those of the existing γS-crystallins characterized from mammalian species. The results indicate that γS crystallin characterized from the lower class of cartilaginous fish such as shark, in contrast to those γS-crystallins of

[1] The sequence data of cDNAs for shark γS-crystallins have been deposited in the EMBL Data Library under the Accession Numbers X79226 and X79227 for γS1 and γS2, respectively.

[2] Corresponding address: Institute of Biochemical Sciences, National Taiwan University, P.O. Box 23-106, Taipei, Taiwan. Fax: (886)-2-3635038.

0006-291X/97 $25.00

A

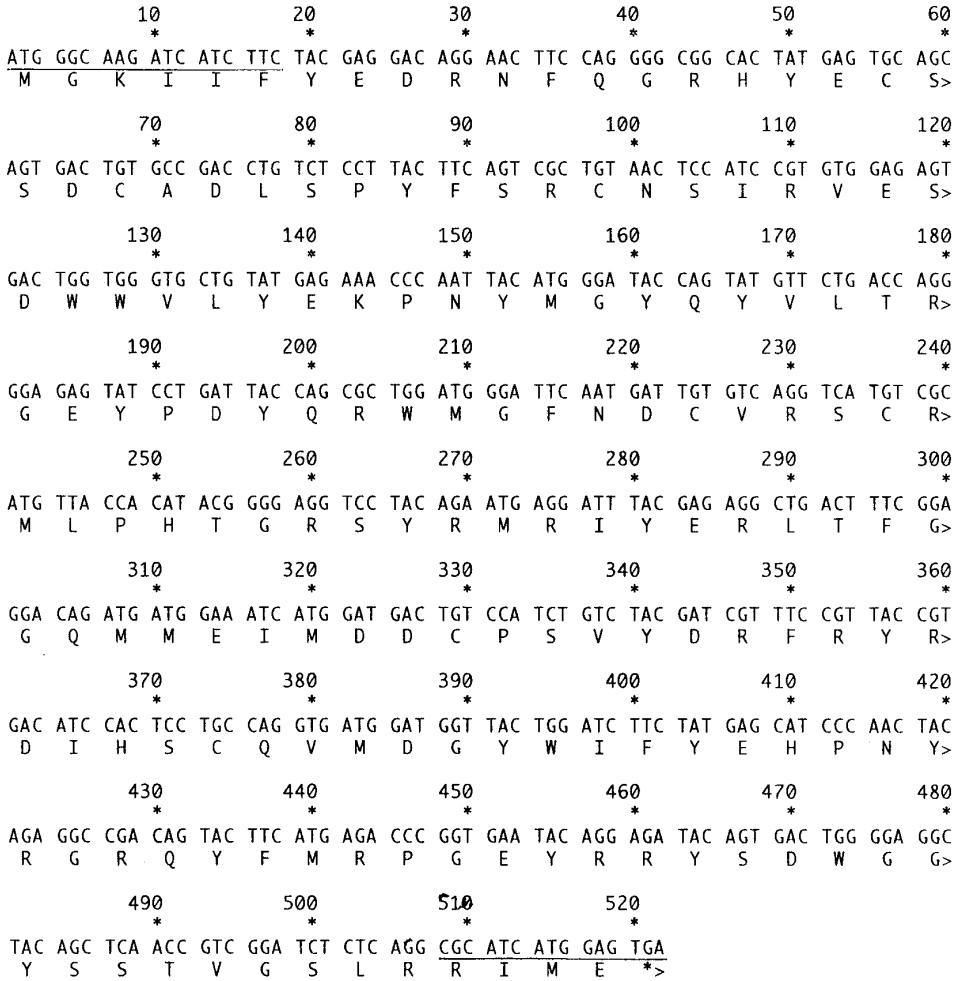

FIG. 1. Nucleotide and deduced protein sequences of shark γS-1 (**A**) and shark γS-2 (**B**) crystallins. In (**A**) the nucleotide sequence of 522-base pairs is shown above the amino-acid sequence of 173 residues including translation initiation methionine; in (**B**) the nucleotide sequence comprises 524-base pairs encoding a protein sequence of 173 amino acids. Asterisks (*) are indicated in every 10-nucleotide segment for easy tracing of sequence contents. Amino acids are denoted by one-letter symbols. The 5′ and 3′ nucleotide segments used as primers for PCR reactions are underlined.

mammalian class, consists of a multitude of isoforms, and are more related to γ-crystallin than γS or β-crystallins based on sequence alignment and phylogenetic comparison.

MATERIALS AND METHODS

Isolation of mRNA from shark lenses. The small shark (*Chiloscyllium colax,* brownbanded bambooshark or brown-spotted catshark as commonly called) of less than 1-year-old was obtained from a local aquarium shop under a special contract for scientific research. Shark lenses were removed and stored in liquid-nitrogen container immediately after they were dissected and before the processing for mRNA isolation. Two deep-frozen lenses from one shark were homogenized and RNA was extracted according to the standard cloning manual of Maniatis et al. (9). To obtain a full-length crystallin cDNA, poly(A)$^+$RNA was purified using QuickPrep mRNA preparation kit (Pharmacia, Uppsala, Sweden) and then subjected to the synthesis of cDNA mixture by cDNA Synthesis System/Plus kit (Amersham, England).

cDNA amplification by PCR, cloning and sequencing of γS-crystallin isoforms. Two oligonucleotide primers of sense and antisense orientations, covering 5′- and 3′-nucleotide coding regions for N- and C-terminal 4-6 amino-acid segments of the previously determined cDNA sequence for one carp γS-crystallin (10), with the forward sequence, 5′-CATGGGCAAG(A/G)TCA(T/C)CTT(C/T)-3′ (19-mer) and the reverse sequence, 5′-C(A/G)TCACTCCA(T/C)(G/A)A(T/G)GCG-3′ (17-mer) (with slant lines indicating use of degenerate codons in the primers) were synthesized. The conditions for PCR reactions were similar to the previous reports for cDNA amplification of teleostean lenses (5,6), *i.e.* subjecting to 40 cycles of heat denaturation at 94°C for 2.5 min, annealing the primers to the DNAs at 48°C for 1 min and 20 sec and running DNA chain extension with *Taq* polymerase at 72 °C for 3 min, followed by a final extension at 72 °C for 10 min. Products were treated with Klenow Fragment and T4 polynucleotide kinase, and separated on a 1.2 % agarose gel and electroeluted according to standard procedures. The DNA fragments were subcloned into pUC18 previously digested with SmaI/BAP, and then transformed into *E. coli* strain JM 109. Plasmids purified from positive clones were prepared for nucleotide sequencing by dideoxynucleotide chain-termination method (11). The DNA sequence deter-

B

```
          10          20          30          40          50          60
           *           *           *           *           *           *
ATG GGC AAG ATC ATC TTC TAC GAG GAC AGG AAC TTC CAG GGG CGG AAC TAT GAG TGC AGC
 M   G   K   I   I   F   Y   E   D   R   N   F   Q   G   R   N   Y   E   C   S>

          70          80          90         100         110         120
           *           *           *           *           *           *
AGT GAC TGT GCC GAC CTG TCT CCT TAC TTC AGC CGC TGT AAC TCC ATC CGT GTT GAG AGT
 S   D   C   A   D   L   S   P   Y   F   S   R   C   N   S   I   R   V   E   S>

         130         140         150         160         170         180
           *           *           *           *           *           *
GAC TGG TGG GTG CTG TAT GAG AAA CCC AAT TAC ATG GGA TAC CAG CAT GTT CTG ACC AGG
 D   W   W   V   L   Y   E   K   P   N   Y   M   G   Y   Q   H   V   L   T   R>

         190         200         210         220         230         240
           *           *           *           *           *           *
GGA GAG TAT CCT GAC TAC CAG CGC TGG ATG GGA TTC AAT GAC TGT GTC AGG TCA TGT CGA
 G   E   Y   P   D   Y   Q   R   W   M   G   F   N   D   C   V   R   S   C   R>

         250         260         270         280         290         300
           *           *           *           *           *           *
GTA CCC ACA CAC ACC CAG AGG CCC TAC AGA ATG AGG ATC TAC GAG AGG CCT GAC TTC GGA
 V   P   T   H   T   Q   R   P   Y   R   M   R   I   Y   E   R   P   D   F   G>

         310         320         330         340         350         360
           *           *           *           *           *           *
GGA CAG ATG ATG GAA TTC ATG GAT GTC TGT CCA TCT GTC TAC GAT CGT TTC CGT TAC CGT
 G   Q   M   M   E   F   M   D   V   C   P   S   V   Y   D   R   F   R   Y   R>

         370         380         390         400         410         420
           *           *           *           *           *           *
GAC ATC CAC TCC TCC CAT GTG ATG GGC GCT TAC TGG ATC TTC TAT GAA CAT CCC AAC TAC
 D   I   H   S   S   H   V   M   G   A   Y   W   I   F   Y   E   H   P   N   Y>

         430         440         450         460         470         480
           *           *           *           *           *           *
AGA GGA CGA CTA TAC TTC ATG CGC CCT GGC GAA TAC AGA AGA TAC AGT GAC TGG GGA GGA
 R   G   R   L   Y   F   M   R   P   G   E   Y   R   R   Y   S   D   W   G   G>

         490         500         510         520
           *           *           *           *
TAC AGC TCA ACT ATC GGA TCT TTC AGG CGC ATC ATG GAG TGA CG
 Y   S   S   T   I   G   S   F   R   R   I   M   E   *   R>
```

FIG. 1—*Continued*

mined by conventional isotope-labeled manual method was double-checked by automatic fluorescence-based sequencing of templates amplified by PCR using model 373A DNA sequencing System (Applied Biosystems Inc., CA, USA) with a *Taq* DyeDeoxy terminator cycle sequencing kit (Applied Biosystems Inc.).

Sequence comparison of deduced shark γS-crystallins and homology search. A commercially available software package (DNASTAR Inc., Madison, WI, U.S.A.) was used for the estimation of sequence homology based on percent sequence identity (5).

Construction of a phylogenetic tree for β-, γ- and γS-crystallins of vertebrate species. A software package of LaserGene for the Apple Macintosh computer from DNASTAR, Inc. was used for the estimation of sequence homology based on percent similarity and divergence among different cDNA sequences of β-, γ- and γS-crystallins. Percent divergence is calculated by comparing sequence pairs in relation to the phylogenetic tree. On the other hand the percent similarity is estimated by comparing sequences directly without accounting for phylogenetic relationships. Phylogenetic tree was then constructed using the algorithm of Hein (12) in the MegAlign programs of the package. It is a multiple-sequence alignment program that builds trees as it aligns DNA or protein sequences using a combination of distance matrix and approximate parsimony methods. This method constructs multiple-alignment by imposing restrictions based on evolutionary relatedness of the aligned sequences, which is useful to align highly evolved gene families that have clear evolutionary relationships.

RESULTS AND DISCUSSION

There were sharks in the oceans of earth long before the first animals had begun to colonize the land surface. Their history stretches back for at least 700 million years, a vast period of time as compared with shorter evolutionary history of other vertebrates (3). That sharks were once considered to be the most primitive are now thought to have been relatively specialized regarding their complex biology to be ranked with birds and mammals as highly evolved creatures. In this report the structural characterization of γS-crystallins from shark lenses by PCR and its comparison with those of teleostean and mammalian lenses is of special important to unravel the complex evolutionary history of lens crystallins in the animal kingdom.

FIG. 2. Multiple sequence alignment and comparison of seven crystallin sequences from species of different classes. The identical amino-acid residues among various sequences based on the first one (shark γS-1) were expressed in white letters against black-background blocks. The gaps were introduced for optimal alignment and maximum homology for the sequences. Note that the middle region (residues# 70–130) shows a greater sequence variation than the N- and C-terminal regions among the compared sequences. Amino acid residues are denoted by one-letter symbols.

Characterization of cDNAs Encoding γS-Crystallins from Shark Lenses

Previous studies have indicated the unusual structural characteristics of shark γ-crystallin as compared with those associated with the lenses of teleostean fishes such as carp (13–15). Especially noteworthy is the finding that the amino acid compositions of γ-crystallins seem to lack the unique characteristic of high methionine content (> 10%) as commonly observed for that of teleostean fishes (5,6). In contrast it is closer to mammalian γ-crystallins regarding both the amino acid composition, N-terminal partial sequence plus some conformational properties as revealed by circular dichroism (14,15). Shark γ-crystallin showed a much more complex pattern in isoelectric focusing (data not shown), revealing that it consists of various charge isoforms (13, 14). Due to the complexity of γ-crystallin, we suspect that the same multiplicity of isoforms may be also present for γS-crystallin, a lens protein with dual structural properties of β- and γ-crystallins (7,8 and the references therein). We have hence resorted to the recent rapid method of cloning and sequencing by means of PCR technique for the determination of cDNA sequences of these multiple isoforms.

PCR amplification of total lens cDNA mixtures prepared from two lenses of a single shark with the designed primers based on partial DNA coding sequences of carp γS-crystallin (10) achieved the isolation of one PCR fragment corresponding to the complete reading frame encoding at least two γS-crystallin isoforms from this shark species. The size determination of PCR-amplified cDNA coding for γS crystallin was estimated to be about 520 bp, in agreement with a protein of about 170–180 amino-acid residues for mammalian γ- and γS-crystallins. The PCR-amplified DNA fragments were then subcloned into pUC18 previously digested with SmaI/BAP, and then transformed into *E. coli*

% homology	Bovine γII	Carp γS	Bovine γS	Bullfrog β	Bovine β	Human γ5
Shark γS-1	64.0	60.0	56.0	35.0	38.0	62.0
Shark γS-2	62.0	58.0	52.0	35.0	37.0	60.0

FIG. 3. Pair-wise comparison of amino-acid sequence homology between two shark γS-crystallins and various β-, γ- and γS-crystallins from species of different classes. Analysis of sequence homology was carried out using the software package (DNASTAR Inc., Madison, WI, USA) on the published sequences of carp γS (10), bovine γS (7), bullfrog β2 (24), bovine β2 (25), bovine γII (26) and human γ5 (27) crystallins.

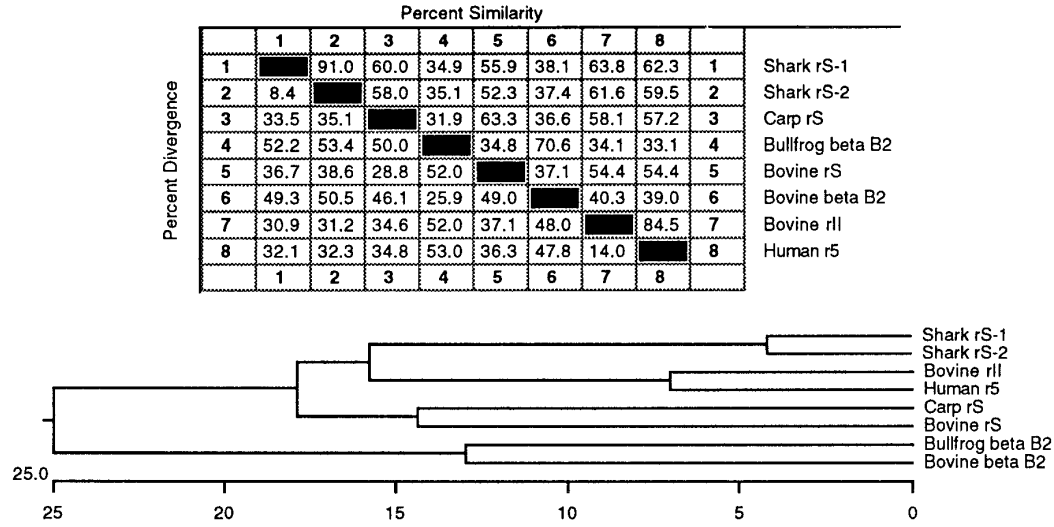

FIG. 4. Pair-wise comparison of protein sequence similarity and divergence (Top) and construction of phylogenetic tree (Bottom) of eight crystallin sequences from various species of three major classes of vertebrates. Analysis of sequence data was carried out in a software package of LaserGene for the Apple Macintosh computer (DNASTAR Inc., Madison, WI, U.S.A.). Percent divergence is calculated by comparing sequence pairs in relation to the relative positions in the phylogenetic tree. On the other hand the percent similarity is estimated by comparing percent sequence identity directly without accounting for phylogenetic relationships. A phylogenetic tree was then constructed based on the percent divergence between protein sequences using a combination of distance matrix and approximate parsimony methods in the phylogeny generation program of Hein (12). This algorithm carries out multiple-alignment by imposing restrictions based on evolutionary relatedness of the aligned sequence pairs. The tree was built using clustal method and weighted residue-weight table. The length of each pair of branches represents the sequence distance between aligned pairs. The scale beneath the tree measures the distance between sequences (in millions of years). Bovine and bullfrog βB2 crystallins are included to indicate the fact that β- and γ-crystallins form a β/γ superfamily.

strain JM 109. Plasmids purified from two positive clones were then prepared for nucleotide sequencing.

Sequence Analysis of cDNA Encoding Shark γS-Crystallins

More than five positive clones have been identified, with their 5′ and 3′ nucleotide sequences being determined to be essentially identical to those of the designed primers, indicative of the existence of multiple isoforms for shark γS-crystallin, which is in contrast to bovine (7,8) and human γS-crystallins with only one sequence being identified (16). In order to avoid sequencing errors introduced in the manual sequencing of isotope-labeled cDNA, sequence accuracy was doubly checked and confirmed by automatic fluorescence-based sequencing technique. The only uncertainty may lie in the first and last few nucleotides present in the 5′ and 3′ region of the PCR fragment even though we have used some degenerate codons in the primers. Further genomic analysis of shark γS-crystallins my help solve the ambiguity in these short segments. The deduced protein sequences together with their genetic coding sequences of two clones, designated as shark γS-1 and γS-2 are shown in **Fig. 1A** and **1B.** The cDNA sequences encoding shark γS-1 and γS-2 were both found to consist of 519 nucleotides excluding the stop codons, each of which covering a full-length protein of 173 amino-acid residues including the initiating methionine, which is close to carp γS (174 a.a.) and slightly lower than bovine γS (178 a.a.).

Sequence Alignment and Comparison of β-, γ- and γS-Crystallins from Different Species

Fig. 2 aligns seven sequences encompassing representative β-, γ- and γS-crystallins from characterized crystallins of the major classes in vertebrates, which have all been deduced from nucleotide sequences coding for these crystallins. It is noteworthy that there is only about 35–64% sequence homology between shark γS crystallins and structurally related β-, γ- and γS-crystallins from different evolutionary classes. However one salient feature is that some of the key residues (such as Tyr-6, Glu-7, Phe-11, Gly-13 and Ser-34 based on bovine γII sequence numbering) for the maintenance of stability in γ-crystallins (17–19) are mostly retained and conserved in all β-, γ- and γS-sequences even from species of distantly related classes, attesting to the conservative structural aspects of β/γ superfamily. It is also of interest to find that N- and C-terminal regions of these crystallins are more conserved than the middle regions of the sequences (residues# 70–130).

In the pair-wise sequence homology comparison (**Fig. 3**) of various deduced amino-acid sequences from species of different classes using commercial software

package (DNASTAR program), it is found that shark γS-crystallins show 62–64% sequence identity to bovine γII crystallins whereas only 52–56% sequence identity is found between shark and bovine γS-crystallins, indicating that shark γS is evolutionarily more related to authentic mammalian γ- than γS-crystallins and may represent one intermediate crystallin form from the divergent evolution of γ-crystallin gene family. On the other hand, shark γS-crystallins show only 58-60% sequence identity to carp γS, underlying the distinct differences of γS-crystallins present in the cartilaginous and teleostean fishes.

Construction of a Phylogenetic Tree Based on Protein Sequence Comparison

We have previously constructed a phylogenetic tree of crystallins based on protein or DNA sequence comparison with similar results (20–23). In **Fig. 4** a phylogenetic tree is constructed based on multiple sequence alignment of eight protein sequences using a combination of distance matrix and approximate parsimony methods (12). It has been shown to be useful to align highly evolved gene families and their corresponding protein sequences that have clear evolutionary relationships as in the cases of various crystallin family. It is noteworthy that the phylogenetic tree based on the sequence divergence among these protein sequences indeed exemplifies the close relatedness between shark γS-crystallins and γ-crystallins from bovine and human lenses. On the other hand, carp γS-crystallin was grouped with bovine γS-crystallin, in agreement with the percent homology shown in Fig. 3. Especially interesting is the observation that two β2-crystallin sequences from bullfrog and bovine indeed locate at different branching points of the tree from those of γ- and γS-crystallins, indicative of two distinct evolutionary pathways leading to β- and γ/γS crystallins from the ancestral β/γ protein.

ACKNOWLEDGMENTS

This work was supported by Academia Sinica and the National Science Council (NSC Grants 83-0203-B-001-086, 83-0418-B-001-020BA, 84-2311-B-001-050-BA & 86-2311-B-002-031-B15), Taipei, Taiwan.

REFERENCES

1. Chiou, S.-H. (1986) *FEBS Lett.* **201,** 69–73.
2. de Jong, W. W., and Hendriks, W. (1986) *J. Mol. Evol.* **24,** 121–129.
3. Keeton, W. T. (1972) *in* Biological Science, 2nd ed., W. W. Norton & Company, New York.
4. Chang, T., Jiang, Y.-J., Chiou, S.-H., and Chang, W.-C. (1988) *Biochim. Biophys. Acta* **951,** 226–229.
5. Pan, F.-M., Chang, W.-C., Chao, Y.-K., and Chiou, S.-H. (1994) *Biochem. Biophys. Res. Commun.* **202,** 527–534.
6. Pan, F.-M., Chang, W.-C., Lin, C.-H., Hsu, A.-L., and Chiou, S.-H. (1995) *Biochem. Mol. Biol. International* **35,** 725–732.
7. Quax-Jeuken, Y., Driessen, H. P. C., Leunissen, J., Quax, W., de Jong, W., and Bloemendal, H. (1985) *EMBO J.* **4,** 2597–2602.
8. van Rens, G. L. M., Raats, J. M. H., Driessen, H. P. C., Oldenburg, M., Wijnen, J. T., Khan, P. M., de Jong, W. W., and Bloemendal, H. (1989) *Gene* **78,** 225–233.
9. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1989) *in* Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
10. Chang, T., and Chang, W.-C. (1987) *Biochim. Biophys. Acta* **910,** 89–92.
11. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
12. Hein, J. (1990) *Methods Enzymol.* **183,** 626–645.
13. Siezen, R. J., Hom, C., Kaplan, E. D., Thomson, J. A., and Benedek, G. B. (1988) *Exp. Eye Res.* **46,** 81–93.
14. Chiou, S.-H. (1989) *FEBS Lett.* **250,** 25–29.
15. Chiou, S.-H., Chen, S.-W., Itoh, T., Kaji, H., and Samejima, T. (1990) *FEBS Lett.* **275,** 111–113.
16. Smith, J. B., Yang, Z., Lin, P., Zaidi, Z., Abbasi, A., and Russell, P. (1995) *Biochem. J.* **307,** 407–410.
17. Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B., and Wistow, G. (1981) *Nature* **289,** 771–777.
18. Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P., and Blundell, T. (1983) *J. Mol. Biol.* **170,** 175–202.
19. Liaw, Y.-C., Chiou, S.-H., Lin, C.-W., Chang, T., and Chang, W.-C. (1992) *J. Biochem.* **112,** 341–345.
20. Chiou, S.-H., Yu, C.-W., Lin, C.-W., Pan, F.-M., Lu, S.-F., Lee, H.-J., and Chang, G.-G. (1995) *Biochem. J.* **309,** 793–800.
21. Lu, S.-F., Pan, F.-M., and Chiou, S.-H. (1995) *Biochem. Biophys. Res. Commun.* **216,** 881–891.
22. Pan, F.-M., Chang, W.-C., Lu, S.-F., Hsu, A.-L., and Chiou, S.-H. (1995) *Biochem. Biophys. Res. Commun.* **217,** 940–949.
23. Lu, S.-F., Pan, F.-M., and Chiou, S.-H. (1996) *J. Protein Chem.* **15,** 103–113.
24. Pan, F.-M., Chang, W.-C., Lu, S.-F., Hsu, A.-L., and Chiou, S.-H. (1995) *Biochem. Biophys. Res. Commun.* **217,** 940–949.
25. Hogg, D., Gorin, M. B., Heinzmann, C., Zollman, S., Mohandas, T. K., Klisak, I. J., Sparkes, R. S., Breitman, M. L., Tsui, L.-C., and Horwitz, J. (1987) *Curr. Eye Res.* **6,** 1335–1341.
26. Bhat, S. P., and Spector, A. (1984) *DNA* **3,** 287–295.
27. Meakin, S. O., Du, R. P., Tsui, L.-C., and Breitman, M. L. (1987) *Mol. Cell. Biol.* **7,** 2671–2679.