Vol. 202, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1994 Pages 527-534

CHARACTERIZATION OF γ-CRYSTALLINS FROM A HYBRID TELEOSTEAN FISH: MULTIPLICITY OF ISOFORMS AS REVEALED BY cDNA SEQUENCE ANALYSIS[1]

F.-M. Pan, W.-C. Chang, Y-.K. Chao and S.-H. Chiou*

Laboratory of Crystallin Research, Institute of Biochemical Sciences, National Taiwan University and Institute of Biological Chemistry, P.O. Box 23-106, Academia Sinica, Taipei, TAIWAN

Received May 24, 1994

γ-Crystallins were isolated and characterized from the eye lenses of a hybrid species belonging to the teleostean fish. Isoelectric focusing of γ-crystallin fraction obtained from gel-permeation chromatography revealed that it consists of multiple charge isomers of a protein species with a molecular mass of about 20 kDa. To facilitate the cloning of γ-crystallin gene, cDNA was constructed from the poly(A)+mRNA isolated from fresh lenses, and amplification by polymerase chain reaction (PCR) was carried out to obtain cDNA encoding multiple γ-crystallins. Sequencing five of more than 10 positive clones revealed that a multiplicity of isoforms exists in the γ-crystallin class of teleostean lenses. Comparison of protein sequences encoded by these multiple cDNAs with those published sequences of γ-crystallins from bovine, mouse and carp lenses indicated that there is about 70-80 % sequence homology between different species of piscine species whereas only 50-60 % is found between mammals and fishes. Structural analysis of these γ-crystallins with high methionine contents (11-16 %) suggests that there are two major subclasses of piscine γ-crystallins, *i.e.* γM_1 and γM_2, existed long before the appearance of mammalian γ-crystallin with low methionines.

Understanding the evolutionary conundrum of various classes of proteins remains a major aim of current research in protein chemistry and molecular biology. The abundant presence of various common and specific classes of structural proteins, *i.e.* lens crystallins, in different species of vertebrates may provide a good model system to unravel the complex process of evolution in structurally homologous proteins [1,2]. More than 90 % of the proteins in lenses of various animal species consist of crystallins of several classes, which exist as water-soluble proteins and were previously thought to possess only the structural role of maintaining lens transparency and optical clarity in the lens fiber cells. Crystallins in three major families, *i.e.*

[1]The sequence data of cDNAs for γ-crystallins reported here have been deposited in the EMBL Data Library under the accession numbers X76322, X76323, X76324, X76325 and X76326 for $\gamma M_{1\text{-}1}$, $\gamma M_{1\text{-}2}$, $\gamma M_{2\text{-}1}$, $\gamma M_{2\text{-}2}$ and $\gamma M_{2\text{-}3}$, respectively.

*To whom correspondence should be addressed.

0006-291X/94 $5.00

those classified as α-, β-, and γ-crystallins, are found almost in all vertebrates [3-5], and other taxon-specific crystallins with enzymatic functions, *e.g.* δ-, ε-, λ- and τ-crystallins, etc., have recently been found to be present in specific species of various classes of animals [6].

Fish represents the oldest and most diverse group of vertebrates. The modern fishes comprise two major classes: (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 would be of special interest from the evolutionary point of view because it constituted the early forms of vertebrates and is thought to have been ancestral to the land vertebrates. The present study was performed in the endeavor to make a systematic characterization of lens crystallins from a teleostean fish using PCR methodology to aid in the structural analysis of multiple isoforms of γ-crystallins. We have isolated and sequenced five clones encoding major γ-crystallins of piscine lenses with high methionine contents (> 11 %). These crystallins constitute the most unusual amino acid composition as compared with the reported γ-crystallins from higher classes of vertebrates and complement our previous sequence characterization of closely-related carp γ-crystallin genes [7,8].

MATERIALS AND METHODS

Isolation and characterization of piscine crystallins

The teleostean fish used for this study is a favored cultivated fish species for pet-fish collector (beautifully red in body color) in Taiwan. It is a hybrid bony fish between ***Petenia splendida*** (female) and ***Cichlasoma synspilum*** (male). Since it is a hybrid by nature, therefore it can not reproduce like its parental generations. They were provided by a local aquarium shop under a special contract for scientific research.

The isolation of lens crystallins from the supernatant of 15,000 x g centrifugation of lens homogenate and the subsequent separation on TSK-HW (55) were as described before [9,10]. The peak fractions were collected and the identification of each class of crystallin was based on their characteristic patterns on SDS-polyacrylamide slab gel (SDS-PAGE).

SDS-polyacrylamide gel electrophoresis (SDS-PAGE, 5 % stacking/ 14 % resolving gel) was as described [11] with some modifications. Isoelectric focusing in 7.5 % polyacrylamide gel containing 2.8 % (v/v) carrier ampholytes of pH 3.5-10 was carried out on a slab gel with the incorporation of 0.1 % 2-mercaptoethanol and 6 M urea [12].

Amino acid analysis and N-terminal sequence analysis by automated Edman degradation with a pulsed-liquid phase sequencer (Model 477A, Applied Biosystems, Foster City, CA. U.S.A.) of lyophilized γ-crystallin fraction from gel-filtration column were carried out as described before [13,14].

Preparation of mRNA from lenses and cloning by PCR amplification

Fish 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 fish were homogenized and RNA was extracted according to the standard procedures [15]. 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).

Three oligonucleotide primers of sense and antisense orientations, covering 5'- and 3'-nucleotide coding regions for N- and C-terminal 5-7 amino-acid segments of two previously determined carp γ-crystallins, *i. e.* γ-m1 and γ-m2, with the forward sequence, 5'-CATGGGCAAG(A/G)TCA(T/C)CTT(C/T)-3' for both γ-m1 and γ-m2, and two reverse sequences, 5'-(T/C)TAACA(G/C)ATATC(A/C)(G/A)TGA(T/C)ACG-3' for γ-m1 and the other 5'-CTAGTACCA(G/C)GA(G/A)TCCATGA(T/C)ACG-3' for γ-m2 (with slant lines indicating use of degenerate codons in the primers) [7], were synthesized. The PCR reactions were subjected to 40 cycles of heat denaturation at 94 °C for 1.5 min, annealing the primers to the DNAs at 48 °C for 2 min 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. The PCR 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 [16].

Sequence comparison of γ-crystallins and homology search

In the comparison and analysis of the deduced amino-acid sequences from determined cDNA sequences coding for γ-crystallins, a software package (DNASTAR Inc., Madison, WI, U.S.A.) was used for the estimation of sequence homology based on percent sequence identity.

RESULTS AND DISCUSSION

Most previous studies on the characterization of crystallins emphasized more on the species of higher vertebrates with relatively fewer reports on the lenses from the lower aquatic vertebrates, *i.e.* various classes of fishes [3]. In this report the characterization of major γ-crystallin class from piscine lenses is of special interest and import in our systematic study of evolutionary relationship of lens crystallins from vertebrates and invertebrates [1,17,18].

Characterization of multiple γ-crystallins from a teleostean fish

Fig. 1 shows the high-resolution isoelectric focusing gel analysis of total lens extract and the crude γ-crystallin fraction isolated from TSK gel permeation column. The identification of γ-crystallin fraction from the column was initially based on its elution position from the column with a native molecular mass of 20 kDa coupled with a similar subunit mass as determined by SDS-gel electrophoresis (data not shown). It is evident that γ-crystallin possesses multiple charge isomeric forms of at least 10 protein species with pI spreading in a range of 6.5 to about

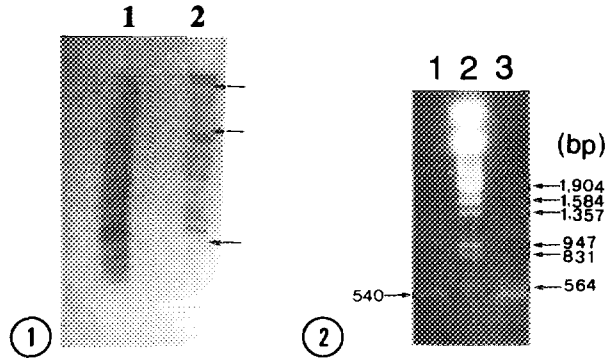

Fig. 1. Isoelectric-focusing gel analysis of fish crystallins from a hybrid teleostean species under denaturing conditions. About 10 μg of total lens extract (lane 1) and unfractionated crude γ-crystallin (lane 2) fractions were layered on a 7.5 % polyacrylamide gel containing 6 M urea and 0.1 % 2-mercaptoethanol for isoelectric focusing (basic end is at the top). The pI was estimated from a pI calibration kit (3.6-10.2) with arrows indicating three marker proteins of pI 5.9, 8.6 and 9.7, respectively.

Fig. 2. Identification and size determination of PCR-amplified γM_1 and γM_2 crystallin cDNA. Electrophoresis was carried out in 1.2 % agarose gel. Lane 1, PCR reaction product of about 540 bp (arrow) coding for γM_1 crystallin; Lane 2, DNA molecular size markers of 0.15 to 2.2 kb (Boehringer Mannheim, Germany). Six major bands ranging from 564 to 1904 bp are indicated by arrows; Lane 3, PCR reaction product of about 540 bp coding for γM_2 crystallin.

10. Similarly, the previous study of shark γ-crystallin also showed complex multiple banding pattern under the same conditions of isoelectric focusing [19]. N-Terminal sequence analysis of the crude γ-crystallin fraction for the first 20 amino acids has also indicated several heterogeneous sites with more than two amino acids being identified along the chain despite the fact that the sequence is in general very similar to that of the carp γ-crystallin [7,13]. Due to the complexity of the γ-crystallin, we have resorted to the recent powerful cloning and sequencing tool of PCR technique for the facile determination of protein sequences of these multiple isoforms by PCR amplification (*vide infra*).

cDNA amplification by PCR and sequence analysis

PCR amplification of total lens cDNA mixtures with the designed primers based on γ-m1 and γ-m2 of carp γ-crystallins [7] achieved the isolation of two PCR fragments corresponding to complete reading frames encoding two homologous γ-crystallin isoforms from this novel species, *i.e.* γM_1 and γM_2 in this report. **Fig. 2** shows the size determination of PCR-amplified cDNA coding for γM_1 and γM_2. The DNA bands were estimated to be about 540 bp, in agreement with a protein of about 170-180 amino-acid residues. The PCR-amplified 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 then prepared for nucleotide sequencing. It is noteworthy that more than 10 positive clones have been identified, with their 5' coding nucleotide sequences being determined to be essentially identical to that of the designed 5' primer. The deduced protein sequences together with their genetic coding sequences of two of the five clones, designated as $\gamma M_{1\text{-}1}$ and $\gamma M_{2\text{-}1}$, are shown in **Fig 3.** The cDNA sequences encoding $\gamma M_{1\text{-}1}$ and $\gamma M_{2\text{-}1}$ were found to consist of 537 and 528 nucleotides respectively, which cover a full-length protein of 177 and 174 amino acids excluding the initiating methionine. We have taken these two as exemplars to illustrate the general sequence characteristics of γM_1 and γM_2 subclasses of piscine γ-crystallin.

Structural comparison of piscine and mammalian γ-crystallins

In the analysis of the deduced amino-acid sequences using commercial software package for sequence analysis (DNASTAR program), it is found that γM_1 and γM_2 subclasses of this hybrid fish show 55-59 % and 47-56 % sequence identity to bovine γII and mouse γ_1/γ_2 crystallins respectively. On the other hand, they show about 70-82 % sequence identity to carp γM_1 and γM_2 crystallins. This certainly underlies the closer evolutionary relationship between these two piscine species as compared with that between fishes and mammals (**Fig. 4**). We have also included bovine and carp γs crystallins [20] (previously called βs crystallins) in the pair-wise comparison with $\gamma M_{1\text{-}5}$. Both bovine and carp γs-crystallins with low methionine seems to share lower sequence homology (< 50 %) with $\gamma M_{1\text{-}5}$ crystallins than that between bovine γII crystallin to $\gamma M_{1\text{-}5}$ crystallins (55-59 %). This seems to imply that in the evolution of γ-crystallin from the piscine γ-crystallins to modern mammalian γ-crystallins, γs-crystallin class may follow a different path from that of mammalian γ-crystallins (bovine γII and mouse γ_1/γ_2) with low methionine content. All these determined sequences have laid a firm molecular basis for the future construction of a molecular phylogenetic tree among these apparently quite heterogeneous and complex families of structurally homologous crystallins .

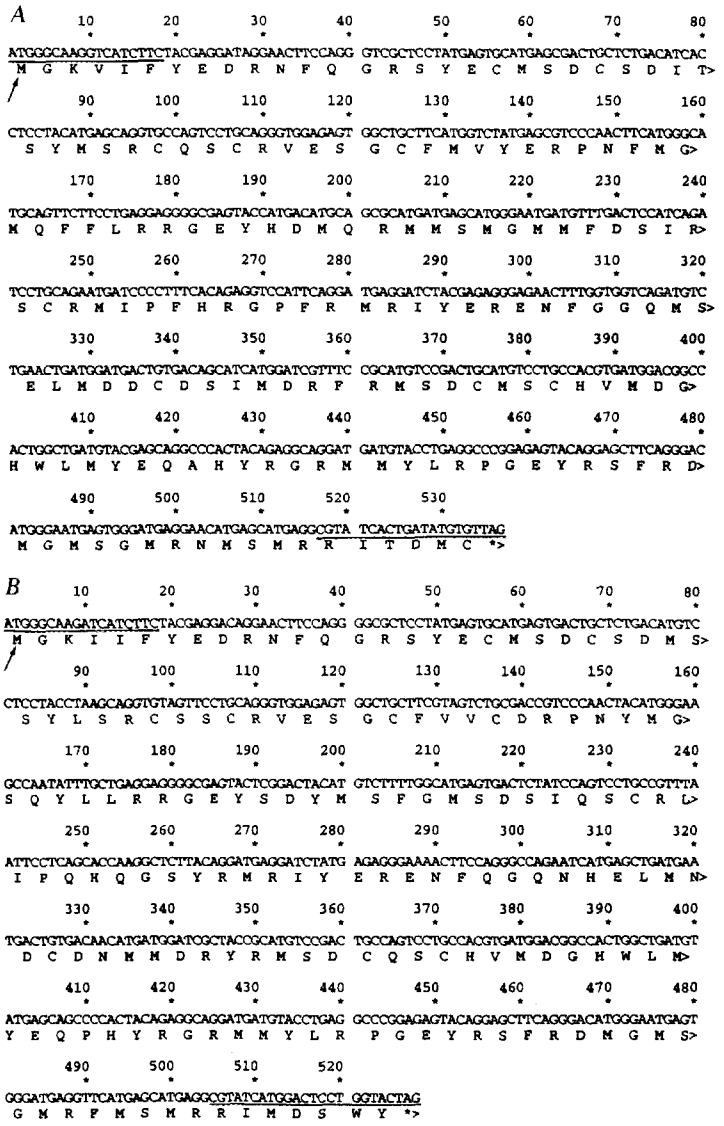

Fig. 3. Nucleotide and deduced protein sequences of $\gamma M_{1\text{-}1}$ (A) and $\gamma M_{2\text{-}1}$ (B) crystallin cDNAs. In (A) the nucleotide sequence of 537-base pairs is shown above the amino-acid sequence of 177 residues excluding translation initiation methionine (arrow) whereas in (B) the nucleotide sequence comprises 528-base pairs encoding a protein sequence of 174 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.

Fig. 5 aligns five determined sequences of $\gamma M_{1\text{-}1}$, $\gamma M_{1\text{-}2}$, $\gamma M_{2\text{-}1}$, $\gamma M_{2\text{-}2}$ and $\gamma M_{2\text{-}3}$, which have all been deduced from cDNAs coding for these γ-crystallins. It is noteworthy that the extent of protein sequence similarity is higher between $\gamma M_{1\text{-}1}$ and $\gamma M_{1\text{-}2}$ (97.0 %) than that of $\gamma M_{1\text{-}1}/\gamma M_{2\text{-}1}$ (79 %) or $\gamma M_{1\text{-}2}/\gamma M_{2\text{-}1}$ (78 %). The % sequence identity among the three γM_2 crystallins is more variable with % identity between each pair ranging from 73-82 %. One

% homology	Bovine γII	Bovine γS	Mouse γ1	Mouse γ2	Carp γM1	Carp γM2	Carp γS
$\gamma M_{1\text{-}1}$	57.10	44.60	51.40	56.50	81.50	73.60	49.20
$\gamma M_{1\text{-}2}$	56.20	43.80	50.60	56.20	81.00	70.90	48.30
$\gamma M_{2\text{-}1}$	54.90	43.60	50.30	55.50	76.10	75.30	48.90
$\gamma M_{2\text{-}2}$	54.90	43.00	47.40	53.80	76.70	78.20	46.60
$\gamma M_{2\text{-}3}$	58.50	41.70	50.00	55.10	70.50	68.80	45.80

Fig. 4. Pair-wise comparison of amino-acid sequence homology among various piscine and mammalian γ-crystallins. Analysis of sequence homology was carried out in the software package (DNASTAR Inc., Madison, WI, USA) using the published sequences for mouse crystallins [22] and various bovine and carp γ-crystallins cited in [23].

salient difference found between γM_1 and γM_2 is at the C-terminal regions, with ITDMC and IMDSWY segments being identified for γM_1 and γM_2 respectively. Therefore judging by the above distinct sequence characteristics, it seems reasonable to conclude that these are two major subclasses of fish γ-crystallins with high methionine contents in contrast to another major mammalian γ-crystallin class, such as bovine γI-γIV [21] and mouse γ-crystallins [22]

```
      *** ** ********** ***    ** *** ********* * *
γM1-1 MGKVIFYEDRNFQGRSYECMSDCSDITSYMSRCQSCRVESGCFMVYER
γM1-2 MGKIIFSEDRNFQGRSYECMSDCSDITSYMSRCQSCRVESGCFMVYER
γM2-1 MGKIIFYEDRNFQGRSYECMSDCSDMSSYLSRCSSCRVESGCFVVCDR
γM2-2 MGKVIFYEDRNFQGRSYETSSDCADMSSYLSRCHSCRVESGCFMVYDR
γM2-3 MGKIIFYEDRNFQGRSYETSSDCAELTSYLSRCNSCRVESGCFMVYER

       * ** *    **** *             * * *** ** * *  **
γM1-1 PNFMGMQFFLRRGEYHDMQRMM-SMGMMF-DSIRSCRMIPFHRGPFRM
γM1-2 PNFMGMQFFLRRGEYHDMQRMM-SMGMMF-DSIRSCRMIPFHRGPFRM
γM2-1 PNYMGSQYLLRRGEYSDYMSF----GMS--DSIQSCRLIPQHQGSYRM
γM2-2 TNYMGNQFFAWRGEYSDYQRMMR-------DCIRSCRMIPMHRGQFRM
γM2-3 PNYMGHQMLARRGEYPDNQRLM---GMSMSDCIRSCRMIPMHRGQFRM

      ***** ** **    ** *    ** ***** *  *** * ***** *
γM1-1 RIYERENFGGQMSELMDDCDSIMDRFRMSDCMSCHVMDGHWLMYEQAH
γM1-2 RIYERENFGGQMSELMDDCDRIKDRFRMSDCMSCHVMDGHWLMYEQAH
γM2-1 RIYERENFQGQNHELMD-CDNMMDRYRMSDCQSCHVMDGHWLMYEQPH
γM2-2 RIYERENFGGQMYE-MDDCDNMMDRYRMSDCQSCHVMDGHWLMYEQPH
γM2-3 RIYERDNFGGQMYELMDGCENIQDRYRMSDCQSANVMDAHCLMYEQPH

      * *************** ***  **  * *** *
γM1-1 YRGRMMYLRPGEYRSFRDMGMSGMRNMSMRRITDMC
γM1-2 YRGRMMYLRPGEYRSFRDMGMSGMRFMSMRRITDMC
γM2-1 YRGRMMYLRPGEYRSFRDMGMSGMRFMSMRRIMDSWY
γM2-2 Y-GRMMYLRPGEYRSFRDMGMSGMRFMSMRRIMDSWY
γM2-3 YRGRMMYLRPGEYRSFREMGMD-MRIGSIRRIMDSWY
```

Fig. 5. Alignment and sequence comparison of five determined sequences $\gamma M_{1\text{-}1}$, $\gamma M_{1\text{-}2}$, $\gamma M_{2\text{-}1}$, $\gamma M_{2\text{-}2}$ and $\gamma M_{2\text{-}3}$ from a teleostean fish species. The identical amino-acid residues among all five sequences are indicated by asterisks (*), which constitute about 63-64 % sequence identity. Amino acid residues are denoted by one-letter symbols.

with typical low methionines. It should also be emphasized that there is about 63-64 % sequence identity among all five γM_1 and γM_2 sequences (as indicated by the asterisks above these sequences in Fig. 5) despite some distinct differences mentioned above.

CONCLUSION

Understanding the mechanism for the evolution of functionally related proteins from different species remains a general biological problem. The structural and genetic basis for the generation of multiple γ-crystallin isoforms is of significant interest, which provides the motivation to compare the structures of this unique crystallin class from more than one piscine species. The recent rapid development of PCR has indeed provided a crucial tool to enrich specific γ-crystallins so that cDNA sequence analysis may be performed easily without the need for time-consuming library construction and screening. Therefore the extensive characterization from the evolutionarily or developmentally unique animal species such as the bony fish species shown in this report may eventually provide some insight into the phenomenon of species diversification and the accompanying molecular origin of these γ-crystallins. Further genomic analysis of γ-crystallin genes of several piscine species based on information obtained from characterization of their corresponding cDNA sequences should shed some light on the evolution of this multigene γ-crystallin family in the animal kingdom.

ACKNOWLEDGMENTS

Support for this research provided by Academia Sinica and the National Science Council, Taipei, TAIWAN (Grants NSC 83-0203-B-001-086 and NSC 83-0418-B-001-020BA) is gratefully acknowledged.

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] Harding, J.J. and Dilley, K.J. (1976) Exp. Eye Res., 22, 1-73.
[4] Bloemendal, H. (1982) CRC Crit. Rev. Biochem. 12, 1-38.
[5] Bloemendal, H. and de Jong, W.W. (1991) Prog. Nucleic Acid Res. Mol. Biol. 41, 259-281.
[6] Wistow, G. and Piatigorsky, J. (1988) Ann. Rev. Biochem. 57, 479-504.
[7] Chang, T., Jiang, Y.-J., Chiou, S.-H. and Chang, W.-C. (1988) Biochim. Biophys. Acta 951, 226-229.
[8] Chang, T., Lin, C.-L., Chen, P.-H. and Chang, W.-C. (1991) Biochim. Biophys. Acta 1090, 261-264.
[9] Chiou, S.-H., Chang, T., Chang, W.-C., Kuo, J. and Lo, T.-B. (1986) Biochim. Biophys. Acta 871, 324-328.
[10] Chiou, S.-H., Chang, W.-C., Pan, F.-M., Chang, T. and Lo, T.-B. (1987) J. Biochem. 101, 751-759.
[11] Laemmli, U.K. (1970) Nature 227, 680-685.
[12] Chiou, S.-H., Lo, C.-H., Chang, C.-Y., Itoh, T., Kaji, H. and Samejima, T. (1991) Biochem. J. 273, 295-300.
[13] Chiou, S.-H., Chen, S.-W. and Lo, T.-B. (1986) FEBS Lett. 209, 107-110.
[14] Chiou, S.-H., Chang, W.-P., Lo, C.-H. and Chen, S.-W. (1987) FEBS Lett. 221, 134-138.

[15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1989) in *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
[16] Sanger, F., Milklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
[17] Chiou, S.-H. (1984) J. Biochem. 95, 75-82.
[18] Chiou, S.-H. (1988) FEBS Lett. 241, 261-264.
[19] Chiou, S.-H. (1989) FEBS Lett. 250, 25-29.
[20] Chang, T. and Chang, W.-C. (1987) Biochim. Biophys. Acta 910, 89-92.
[21] Chiou, S.-H., Azari, P. and Himmel, M.E. (1988) J. Protein Chem. 7, 67-80.
[22] Goring, D.R., Breitman, M.L. and Tsui, L.-C. (1992) Exp. Eye Res. 54, 785-795.
[23] Liaw, Y.-C., Chiou, S.-H., Lin, C.-W., Chang, T. and Chang, W.-C. (1992) J. Biochem. 112, 341-345.