Biochimica et Biophysica Acta, 951 (1988) 226–229
Elsevier

BBA 90122

BBA Report

Carp gamma-crystallins with high methionine content: cloning and sequencing of the complementary DNA

Tschining Chang, Yun-Jin Jiang, Shyh-Horng Chiou and Wen-Chang Chang

Institute of Biochemical Sciences, National Taiwan University and Institute of Biological Chemistry, Academia Sinica, Taipei (Taiwan, China)

(Received 16 May 1988)
(Revised manuscript received 11 August 1988)

Key words: γ-Crystallin; cDNA; Amino acid sequence; Methionine content; Sequence homology; DNA sequence; (*C. carpio*)

The nucleotide sequences of γ-crystallin cDNAs cloned from the common carp (*Cyprinus carpio*) have been determined. The amino-acid sequences derived consist of two polypeptides with 177 and 172 amino-acid residues for γ-m1 and γ-m2, respectively. They exhibit unusually high methionine contents: 12.4% for γ-m1 and 14% for γ-m2. Comparison of both fish γ-crystallins with bovine γ-II crystallin reveals that they are similar in structure. The striking features of both fish γ-crystallins are as follows. (1) Both of them retain the 'conserved' residues, i.e., Tyr-6, Glu-7, Gly-13, Ser-34 and their equivalents in other motifs. (2) they possess the second aromatic residue at position 11. Both of these structural features are considered to be the major factors in stabilizing the folded hairpin structure of the protein. (3) The variable residues in the core region of C-terminal domain are almost all sulfur-containing amino acids, i.e., methionine or cysteine. (4) 30% of the surface hydrophobic groups are composed of methionine. The last two unusual features have been found so far only in these two fish γ-crystallins. The high methionine content may make an important contribution to the protein stability of fish γ-crystallins.

The eye lenses of vertebrates form a complex conundrum of protein evolution as judged by the abundant presence of various common and specific classes of structural proteins, i.e., lens crystallins, in different species of vertebrates [1]. Lens crystallins comprise several major classes of proteins with various extents of heterogeneity [2,3]. Recent progress in recombinant-DNA techniques has facilitated the elucidation of gene structures and their corresponding protein sequences from several different species [4–7]. The comparison of amino-acid compositions of γ-crystallins from different species reveals a high methionine content in aquatic lens crystallins [8]. In this study, we report the nucleotide sequences of two cDNAs encoding the major γ-crystallin polypeptides of high methionine content. These polypeptides are of unusual amino-acid composition compared to other γ-crystallins from higher vertebrates and they complement our previous sequence characterization of another closely related carp β_s crystallin gene [9].

The cDNA library of carp lenses was constructed as described previously [9] with the ex-

The sequence data in this paper have been submitted to the EMBL/Genbank Data Libraries under the accession numbers X12902 (γ-m1) and X12903 (γ-m2).

Correspondence: W.C. Chang, Institute of Biological Chemistry, Academia Sinica, P.O. Box 23-106, Taipei, Taiwan, Republic of China.

ception of using dG-tailed pUC19 as vector. Positive colonies were selected by the method of Southern blotting [10] with carp β_s [9] fragment (nucleotide residues 56–430) as probe. One of the positive clones, pC18, was sequenced and found to contain 14 ATG codons in frame (data not shown). The pC18 was treated with *Bam*HI and *Hin*fI to cut off the dG · dC homopolymer together with

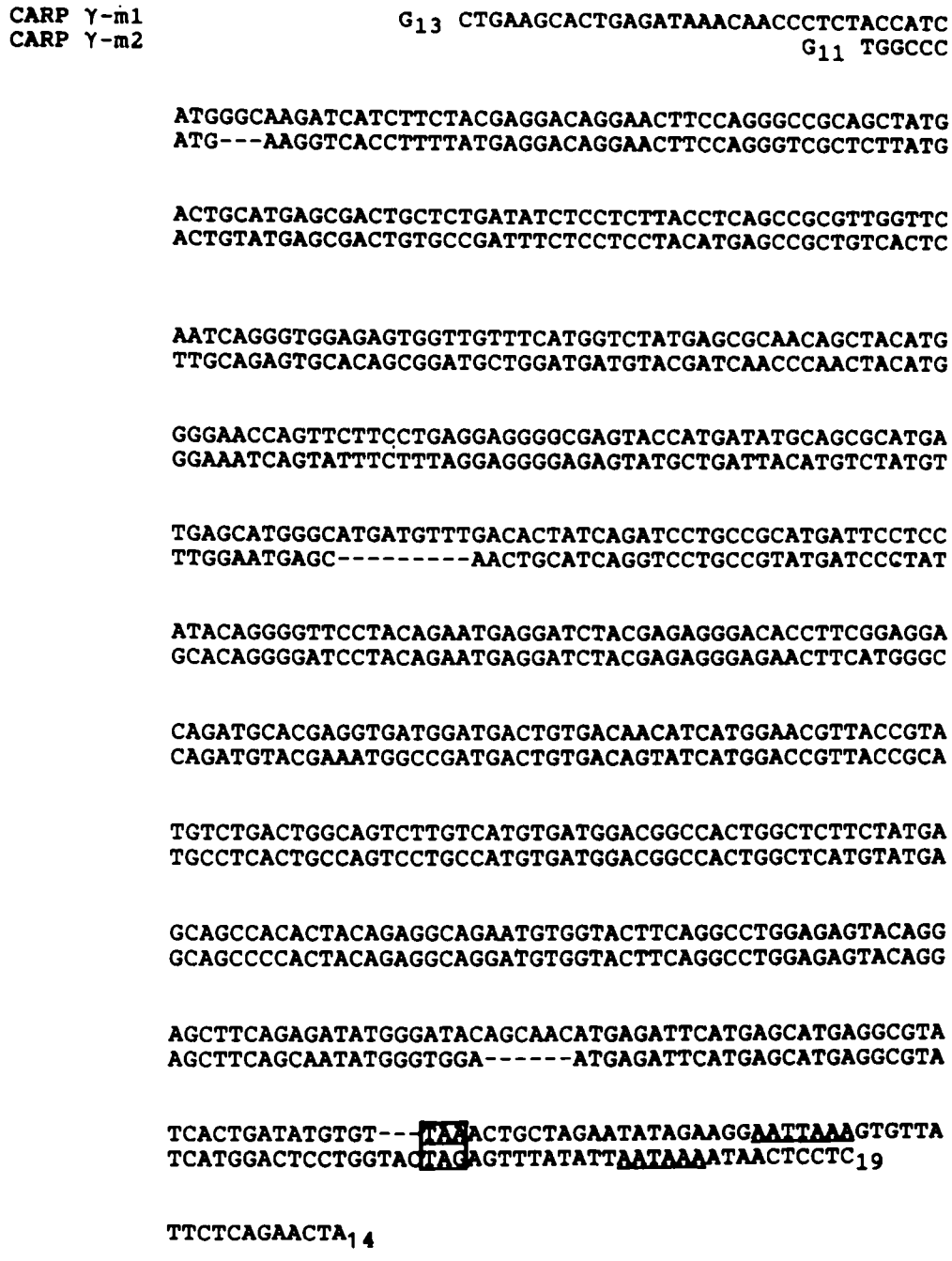

Fig. 1. Comparison of cDNA sequences for carp γ-m1 and γ-m2 crystallins. The sequences are aligned according to their corresponding amino-acid sequence. The stop codons for the longest open reading frame are blocked. Polyadenylation signals for these two cDNAs are underlined.

poly(A) stretches. The resultant ATG-rich DNA fragment was used as probe to re-screen the cDNA library. Two of the positive clones, γ-m1 and γ-m2, which carry full-sized cDNA, were subjected to sequence determination by chemical [11] or supercoil sequencing [12] methods. The complete nucleotide sequence is shown in Fig. 1. Not counting the dG · dC homopolymer, the total length of γ-m1 is 627 base-pairs (bp) and γ-m2 is 550 bp (Fig. 1). The 5′ noncoding region of γ-m1 upstream from the initiation ATG codon consists of 33 nucleotides and the sequences flanking the ATG codon are consistent with the consensus sequence (purine-X-X-A-T-G-G). However, γ-m2 has only 6 nucleotides upstream from the first ATG codon and lacks the consensus sequence. The possibility that γ-m2 is part of a β-crystallin gene cannot be ruled out, since no information about fish β-crystallin is available at the present time.

In the 3′ noncoding region we found the polyadenylation signal, AATTAAA, in γ-m1 which is dissimilar to the common sequence of AATAAA for γ-m2, carp β_s, calf β_s and calf γ-II [9,13,14]. It is noteworthy that at 5′-end of γ-m2 there is a deletion of GGC codon for glycine just after the initiation codon as compared to γ-m1. The derived amino-acid sequences are shown in Fig. 2.

The homology of the coding sequence between these two cDNA is 75%, which is lower than those of other known mammalian γ-crystallins of the same species, i.e., 84% for human [18] and 86% for mouse [16], but similar to that of frog, 69% [17]. The homology of the derived amino-acid sequences is also lower, 73% for carp (Fig. 3), 79% for human [15] and 81% for mouse [16]. If γ-m1 is used as reference for comparison, the sequence of γ-m2 reveals much more deletions than other known mammalian and amphibian γ-crystallin genes. The possibility remains that the carp γ-crystallin gene has follow an evolutionary pathway different from that of higher species, but this hypothesis would be proved only when the genome sequence of fish γ-crystallin gene become available.

The fish γ-crystallins reported here contain the unusually high content of methionine (12.4% for γ-ml and 14% for γ-m2) in amino-acid compositions, similar only to haddock γ-crystallin fraction IV (20.9%) [18] and the invertebrate crystallin of squid lens (12.5%) [8]. Both the N-terminal and C-terminal segments of carp crystallin are almost

```
                    ••   • •                             •
CARP γ-m1    GKIIFYEDRNFQGRSYDCMS DCSDISSYLS         RVGSIR VES
CARP γ-m2     KVTFYEDRNFQGRSYDCMS DCADFSSYMS         RCHSCR VSH        MOTIF 1
CALF γ-II    GKITFYEDRGFQGHCYECSS DCPNLQPYFS         RCNSIR VDS
                      10        20          30

             YRMRIYERDNFGGQMHEVMD DCDNIMERYRMS       DWQSCH VMD
             YRMRIYERENFMGQMYEMAD DCDSIMDRYRMP       HCQSCH VMD        MOTIF 3
             FRMRIYERDDFRGQMSEITD DCPSLQDRFHLT       EVHSLN VLE
               90        100        110              120

             GCFMVYERNSYMGNQFFLRRGEYHDMQR MMSMGMMFDTIRSCRMIP PYRGS
             GCWMMYDQPNYMGNQYFFRRGEYADYMS MFGMS    NCIRSCRMIP MHRGS    MOTIF 2
             GCWMLYERPNYQGHQYFLRRGDYPDYQQ WMGFN    DSTRSCRLIP QHTGT
             40        50        60          70          80

             GHWLFYEQPHYRGRMWYFRPGEYRSFRD MGYSN    MRFMSMRRITDMC
             GHWLMYEQPHYRGRMWYFRPGEYRSFSN MGG--    MRFMSMRRIMDSWY      MOTIF 4
             GSWVLYEMPSYRGRQYLLRPGEYRRYLD WGAMN    AKVGSLRRVMDFY
              130       140       150       160          170
```

Fig. 2. Comparison of the carp γ-m1, γ-m2 and calf γ-II crystallins. Protein sequences are shown in the single-letter amino-acid code. The sequences are displayed by placing motif 1 on top of motif 3 and motif 2 on top of motif 4 so that topologically equivalent amino-acid residues can be easily compared. Solid circles indicate the conserved residues.

the same as that of reported for haddock, as shown below:

carp H_2N-G-K-I-I-F-Y-E-D-; -I-T-D-M-C-COOH

haddock H_2N-G-K-I-T-F-Y-E-D-; -I-T-D-M-C-COOH

However, there is no apparent homology between the N-terminal segments of carp and squid crystallins [19]. It is a prerequisite to obtain the primary sequences of squid crystallin in order to have an unequivocal comparison of phylogenetic relationship between the vertebrate and invertebrate crystallins.

Compared with the amino-acid sequence of calf γ-II, both γ-m1 and γ-m2 were found to retain the conserved residues, i.e., Tyr-6, Glu-7, Gly-13, Ser-34 and their equivalents in the other three motifs. Those residues are important in stabilization of the folded hairpin of γ-crystallins [20,21]. They also possess a second aromatic residue (Phe-11) which packs against the conserved Tyr to stabilize the structure further. In both γ-m1 and γ-m2, these solvent-exposed aromatic and its nearby residues in the tertiary structure are also conserved, i.e., sheet 1: Arg-36, Phe-11; sheet 2: Arg-79, Asp-21, Tyr-50; sheet 3: Arg-152, Glu-150, Phe-98 and sheet 4: Arg-168, Asp-108, Tyr-139 [22]. Since the putative disulfide bond, Cys-18–Cys-22, is also retained, together with those structure maintaining factors mentioned above, both γ-m1 and γ-m2 should assume the similar tertiary structure of calf γ-II crystallin. Both γ-m1 and γ-m2 have a pentapeptide connector between motif 2 and motif 3. They also have a dipeptide extension for γ-m1 and tripeptide extension for γ-m2 at the C-terminus. All of these are unique features of γ-crystallins. Since the γ-m1 and γ-m2 crystallins (Fig. 2) have revealed the similar features of mammalian γ-crystallins, it is very likely that both of them are major components of carp γ-crystallin family.

The most unusual feature of both carp γ-crystallins is the distribution of some hydrophobic residues on the surface and in their core structure. In the core structure, the N-terminal domain is more conserved. But around the variable residues in the C-terminal domain are almost sulfur-containing amino acids: Met-90, Met-105/106, Met-113, Cys-121, Cys-124, Met-165, Met-167. Also, 30% of the surface hydrophobic groups are composed of methionine: Met-68, -69 (γ-m1), -71, -80, -99 (γ-m2), -102, -118, -127. Because the core region is usually composed of hydrophobic residues and aromatic groups, interaction of the polarizable moieties between aromatic groups and methionines might exist. If so, they may contribute significantly to the protein stability [22]. The crystal-structure study of carp γ-crystallin should shed some light on the function of those methionine residues in the protein.

References

1 De Jong, W.W. and Hendriks, W. (1986) J. Mol. Evol. 24, 121–129.
2 Bloemendal, H. (1985) Exp. Eye Res. 41, 429–448.
3 De Jong, W.W. (1981) in Molecular and Cellular Biology of the Eye Lens 9Bloemendal, H., ed.), pp. 221–278, John Wiley & Sons, New York.
4 King, C.R., Shinohara, T. and Piatigorsky, J. (1982) Science 215, 985–987.
5 Yasuda, K., Kondoh, H., Okada, T.S., Nakajima, N. and Shimura, Y. (1982) Nucleic Acids Res. 10, 2879–2891.
6 Tomarev, S.I., Zinovieva, R.D., Chalovka, P., Krayev, A.S., Skryabin, K.G. and Gause, G.G., Jr. (1984) Gene 27, 301–308.
7 Den Dunnen, J.T., Moormann, R.J.M., Lubsen, N.H. and Schoenmakers, J.G.G. (1986) J. Mol. Biol. 189, 37–46.
8 Chiou, S.-H. (1986) FEBS Lett. 201, 69–73.
9 Chang, T. and Chang, W.-C. (1987) Biochim. Biophys. Acta 910, 89–92.
10 Southern, E. (1975) J. Mol. Biol. 98, 503.
11 Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
12 Chen, E.Y. and Seeburg, P.H. (1985) DNA 4, 165–170.
13 Bhat, S.P. and Spector, A. (1984) DNA 3, 287–295.
14 Quax-Jeuken, Y., Driessen, H., Leunissen, J., Quax, W., De Jong, W. and Bloemendal, H. (1985) EMBO J. 4, 2597–2602.
15 Meakin, S.O., Breitman, M.L. and Tsui, L.-C. (1985) Mol. Cell. Biol. 5, 1408–1414.
16 Breitman, M.L., Lok, S., Wistow, G., Piatigorsky, J., Treton, T.A., Gold, R.J.M. and Tsui, L.-C. (1984) Proc. Natl. Acad. Sci. USA 81, 7762–7766.
17 Tomarev, S.I., Zinovievz, R.D., Chalovka, P., Krzyev, A.S., Skryabin, K.G. and Gause, G.G., Jr. (1984) Gene 27, 301–308.
18 Croft, L.R. (1973) Biochim. Biophys. Acta 295, 174–177.
19 Chiou, S.-H., Chang, W.-C., Pan, F.-M., Chang, T. and Lo, T.-B. (1987) J. Biochem. (Tokyo) 101, 751–759.
20 Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P. and Blundell, T. (1983) J. Mol. Biol. 170, 175–202.
21 Summers, L.J., Slingsby, C., Blundell, T.L., Den Dunnen, J.T., Moormann, R.J.M. and Schoenmakers, J.G.G. (1986) Exp. Eye Res. 43, 77–92.
22 Bodner, B.L., Jackman, L.M. and Morgan, R.S. (1977) Biophys. J. 17, 57a.