The nucleotide sequence of a cloned cDNA corresponding to one of the γ-crystallins from the eye lens of the frog *Rana temporaria*

Stanislav I. Tomarev+, Alexander S. Krayev, Konstantin G. Skryabin, Alexander A. Bayev and Georgyi G. Gause jr+

+ *Institute of Developmental Biology* and *Institute of Molecular Biology, USSR Academy of Sciences, Moscow 117334, USSR*

Received 26 July 1982

DNA sequence *γ-Crystallin* *Evolution* *Ophthalmology* *Eye lens* *Protein structure*

1. INTRODUCTION

γ-Crystallins are a family of structurally related proteins present in the eye lens of vertebrates. These proteins are monomeric and of ~20 000 M_r. The amino acid sequences of several bovine γ-crystallins have been published [1–3].

Topologically, they consist of 4 similar folding units (FU) organized into 2 domains [4,5].

γ-Crystallins share sequence homology with at least one β-crystallin; γ- and β-crystallins appear to belong to a single superfamily of proteins [6,7]. Further studies of the structure and evolution of γ-crystallins may benefit from the availability of their cloned genes [8,9].

Here, we report the nucleotide sequence of cDNA corresponding to γ-crystallin from the lens of the frog, *Rana temporaria*.

2. MATERIALS AND METHODS

The construction of cDNA library and the identification of the clone pRt(1) 294 coding for γ-I-crystallin have been described* [9]. Structure determination was performed by the Maxam and Gilbert method after either 3′- or 5′-terminal labeling [10]. More than 90% of the sequence has been determined for both complementary strands.

* For the lack of a commonly accepted nomenclature we shall call γ-crystallin coded by pRt(1) 294 plasmid, γ-I-crystallin

3. RESULTS AND DISCUSSION

The sequencing strategy is shown in fig.1 and the complete nucleotide sequence of γ-I-crystallin cDNA is presented in fig.2. Only one DNA strand corresponding to mRNA is shown. It contains a long open reading frame – 6–399, the derived amino acid sequence shows high homology with that of bovine γ-II-crystallin (see below).

A non-coding region at the 3′-end (position 403–464) has 62 nucleotides in length. The sequence corresponding to the 5′-end of mRNA has been lost in cloning and is substituted by an artefactual sequence (– 41–2 marked by asterisks), which is an inverted repeat of the sequence 439–478 located downstream from the polyadenylation site AATAAA (437–442). The generation of artefactual repeated sequences during cloning has

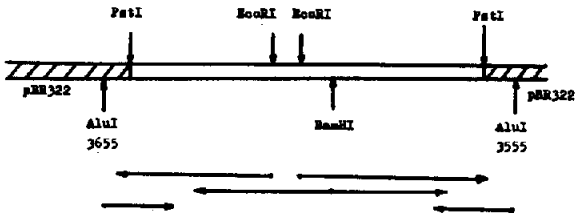

Fig.1. Cleavage map of γ-I-crystallin cDNA region between the 2 *Pst*I sites and the strategy of sequencing. The sequences from the *Bam*HI site were read both from the 3′- and 5′-labeled strands. *Alu*I 3655 and *Alu*I 3555 refer to the coordinates of *Alu*I sites in pBR 322.

Published by Elsevier Biomedical Press
00145793/82/0000–0000/$2.75

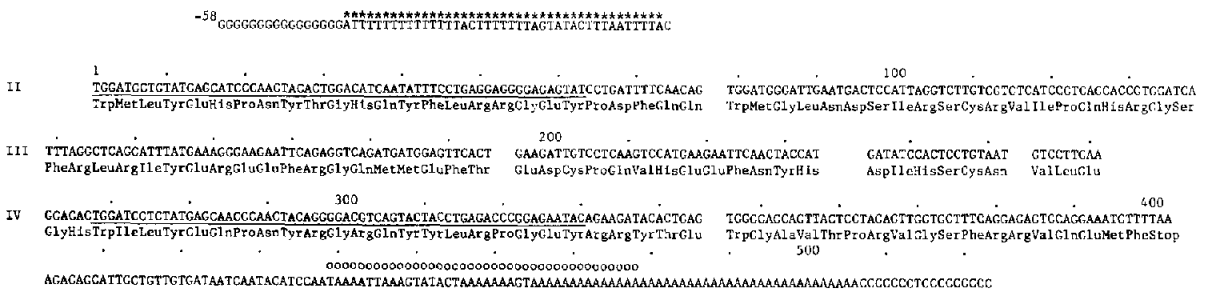

Fig.2. Nucleotide sequence of γ-I-crystallin cDNA. The regions corresponding to γ-I-crystallin folding units II, III and IV are written one under another. Gaps are introduced into the sequence to allow for better alignment between the folding units. The regions of duplication are underlined. The artefactual sequence at the 5′-end is marked with asterisks. The corresponding sequence at the 3′-end is marked with (°) symbols.

been described, and models for these phenomena have been suggested [11,12].

The coding region in the cloned cDNA probably begins from the first letter of tryptophan 41 codon: this amino acid is conservative in at least 2 other members of the β, γ-crystallin superfamily, γ-II-crystallin and βBp crystallin [4,7].

Assuming that frog and bovine γ-crystallins have similar size, the region lost during cloning corresponds to ~40 amino acids and >¾ of the coding sequence of γ-crystallin I mRNA is present in the clone pRt(1) 294. The comparison of cDNA sequences corresponding to the FU II, III and IV reveals considerable homology between FU II and IV and lesser but still significant homology between FU III and the other two. The homology between FU II and IV amounts to 55%, homologies between FU III and II or III and IV amount to 30–35%. Parts of the sequence, however, show markedly greater homology amounting to 76% for regions 1–63 and 268–330 of the FU II and IV. These observations are in agreement with the hypothesis that γ-crystallin sequences originated in evolution by duplications of a hypothetical ancestral gene and that the regions of higher homology correspond to a part of the γ-crystallin structure essential for the interdomain interaction and therefore having greater constraint upon the sequence.

Thanks to the courtesy of Professor Joram Piatigorsky from NIH (USA) we had an opportunity to compare the predicted amino acid sequence of the frog lens γ-I-crystallin derived from cDNA with the Dayhoff's bank of protein sequences. The greatest homology was found as expected with bovine γ-II-crystallin [2], the second best fit was with bovine βBp crystallin [6].

We have conducted a more detailed comparison of the amino acid sequence of the frog lens γ-I-crystallin, bovine γ-II-crystallin and bovine βBp crystallin (fig.3). The sequence of bovine γ-II-crystallin [2] was refined on the basis of X-ray analysis [4]. The homology of frog γ-I-crystallin amino acid sequence with the refined sequence is much better than that with the original sequence. Therefore the

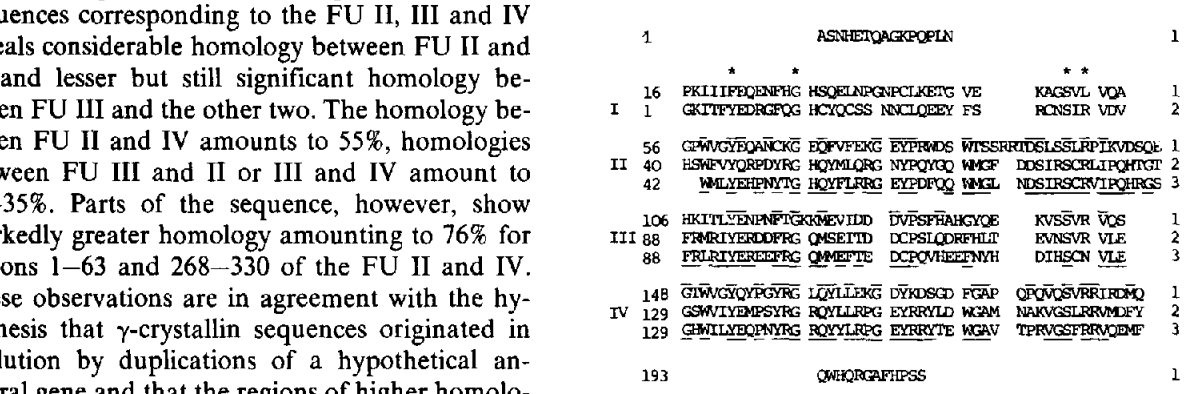

Fig.3. Comparison of the γ-I-crystallin amino acid sequence (3) with amino acid sequence of bovine γ-II-crystallin (2) and βBp-crystallin (1). The sequences are presented as folding units II, III and IV according to [4]. The residues identical in frog γ-I-crystallin and bovine γ-II-crystallin are underlined, the residues identical in frog γ-I-crystallin and βBp-crystallin of the ox are overlined, conservative residues are marked with asterisks.

Protein	Residues	
γ-I-crystallin	91-121	riyErEEFrgqMmEfTedCpqvheefnyhdi
HS 70 kD protein	582-612	ttaEkEEFdhkMeElTrhCspimtkmhqqga
γ-I-crystallin	143-171	qyYLRPgeyrRyteWgavtprvgsfrrvq
αB_2-crystallin	46-74	pfYLRPpsflRapsWidtglsemrlekdr

Fig.4. Homologies between the amino acid sequences of frog lens γ-I-crystallin, *Drosophila* heat-shock 70 000 M_r protein and human/bovine αB_2-crystallin. Identical amino acids in each pair of sequence are typed in capital letters.

subsequent comparison was made with the refined sequence.

The homology between frog lens γ-I-crystallin and bovine γ-II-crystallin is equal to 60% (80 residues out of 133 are identical), the homology is greatest in FU II and IV.

The homology between frog γ-I-crystallin and bovine βBp crystallin is equal to ~35% [7].

The comparison of amino acid sequences of γ-II-crystallin and βBp crystallin has allowed to identify a number of highly conservative amino acid residues in different FU of the molecule [4]. These amino acids are as a rule conserved in the amino acid sequence of frog γ-I-crystallin. They are marked with asterisks over corresponding positions (fig.3). These residues were postulated to be essential for the folding of the γ-crystallin molecule and we assume that the tertiary structure of the frog lens γ-I-crystallin is similar to the tertiary structure of bovine γ-crystallins.

The comparison of the γ-I-crystallin amino acid sequence with Dayhoff's bank has led to the detection of less extensive homologies, some of which may be interesting (fig.4). In this comparison test pieces of the frog lens γ-I-crystallin having 29–31 amino acids in length have been compared with the Dayhoff's bank database. Two interesting homologies should be mentioned:

(i) The region 143–171 has homology with bovine and human αB_2 crystallin residues 46–74; 6 out of 29 residues are identical.

(ii) The sequence 91–121 shows ~26% homology with the region 582–612 of the major 70 000 M_r heat-shock (HS) protein from *Drosophila*.

A partial sequence homology between the sequence of mammalian α-crystallin and 4 small HS proteins of *Drosophila* was described in [13]. This observation together with the finding here of homology between γ-crystallin and 70 000 M_r HS protein of *Drosophila*, raises a possibility that HS proteins and eye lens crystallins might have common evolutionary ancestors. Another possibility is that HS proteins and crystallins interact with similar intracellular targets. Cytoskeleton and chromatin are possible candidates. This hypothesis may serve as a basis of further studies.

REFERENCES

[1] Bjork, I. (1964) Exp. Eye Res. 3, 16–26.
[2] Croft, L.R. (1972) Biochem. J. 128, 961.
[3] Slingsby, C. and Croft, L.R. (1978) Exp. Eye Res. 26, 291.
[4] Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. and Wistow, G. (1981) Nature 289, 771.
[5] Chirgadze, Yu.N., Sergeev, Yu.V., Fomenkova, N.P. and Oreshin, V.D. (1981) FEBS Lett. 131, 81–84.
[6] Driessen, H.P.C., Herbrink, P., Bloemendal, H. and De Jong, W.W. (1980) Exp. Eye Res. 31, 243.
[7] Wistow, G., Slingsby, C., Blundell, T., Driessen, H., De Jong, W.W. and Bloemendal, H. (1981) FEBS Lett. 133, 9–16.
[8] Kozlov, K.A., Tomarev, S.I., Dzhumagaliev, E.B., Dolgilevich, S.M., Skobeleva, N.A., Frolova, L.Yu. and Gause, G.G. jr (1980) Dokl. Acad. Sci. USSR 255, 1503–1506.
[9] Tomarev, S.I., Dolgilevich, S.M., Kozlov, K.A., Zinovieva, R.D., Dzhumagaliev, E.B., Kogan, G.L., Skobeleva, N.A., Mikhailov, A.T., Frolova, L.Yu. and Gause, G.G. jr (1982) Gene 17, 131–138.
[10] Maxam, A.M., Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560.

[11] Weaver, C.A., Gordon, D.F., Kemper, B. (1981) Proc. Natl. Acad. Sci. USA 78, 4073.
[12] Volckaert, G., Tavernier, J., Derynck, R., Devos, R. and Fiers, W. (1981) Gene 15, 215.
[13] Ingolia, T.D. and Craig, E.A. (1982) Proc. Natl. Acad. Sci. USA 79, 2360–2364.