The absence of the long 3′-non-translated region in mRNA coding for eye lens αA_2-crystallin of the frog (*Rana temporaria*)

Stanislav I. Tomarev, Rina D. Zinovieva, Svetlana M. Dolgilevich, Alexander S. Krayev+, Konstantin G. Skryabin+ and Georgyi G. Gause jr

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

Received 1 July 1983

The nucleotide sequence of a cloned cDNA (clone pRt(1)297; GENE (1982) 17, 131) coding for a 18 kDa polypeptide of the frog eye lens has been determined. The sequence, 791 nucleotide in length has only one long open reading frame (447 nucleotides). The derived amino acid sequence in this frame has >90% homology with the region 25–173 of αA_2-crystallin amino acid sequence from a related frog species *Rana pipiens*. The 5′-terminal part of mRNA corresponding to the first 24 amino acids of αA_2-crystallin has been lost in cloning and substituted by an artefactual sequence. The 3′-terminal part appears to be intact as follows from the presence of the universal poly(A) addition site and poly(A) tract. The 3′-non-translated region present in frog αA_2-crystallin mRNA (130 nucleotides) is about 4-times shorter than in mammalian αA_2-crystallin mRNA. Intact αA_2-crystallin mRNA with a size of about 700 nucleotides as determined by Northern blot hybridization is about twice smaller than corresponding mammalian mRNAs.

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

1. INTRODUCTION

Crystallins are structural proteins of the vertebrate eye lens [1]. To gain an insight into the structure of the genes coding for these proteins, we have, in previous papers, described the construction of cDNA recombinant clones corresponding to the nucleotide sequences of total mRNA from the eye lens of the frog *Rana temporaria* [2,3].

In this paper we describe the identity of the recombinant cDNA clone pRt(1)297 from the frog cDNA clonotheque. This clone has been demonstrated to code for an 18 kDa eye lens polypeptide [3]. The identification achieved by DNA sequencing has demonstrated that the cDNA of this clone codes for αA_2 lens crystallin. An unexpected feature of the frog eye lens αA_2-crystallin mRNA sequence was the presence of a 3′-non-translated region (3′-NTR), which had a markedly lower length than in known mammalian αA_2-crystallin mRNA.

2. MATERIALS AND METHODS

The library of cDNA recombinant clones, as well as the clone pRt(1)297, which codes for an 18 kDa polypeptide in a hybrid-selected translation test has been described in [3]. Structure determination was performed by the Maxam and Gilbert method after 3′- or 5′-terminal labelling [4]. More than 95% of the sequence was determined independently for both complementary strands.

Northern blot hybridization was carried out after electrophoretic separation of frog lens poly(A)+-RNA in an agarose formaldehyde gel and the transfer onto nitrocellulose [5]. Nick-translated [^{32}P]DNA of pRt(1)297 plasmid was used as a probe. Computer analysis of the sequences was made, using a version of the dot-matrix procedure [6].

3. RESULTS AND DISCUSSION

The restriction map and sequencing strategy of

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

the cDNA present in pRt(1)297 plasmid is shown in fig.1, and the nucleotide sequence obtained is presented in fig.2. Only one DNA strand corresponding to mRNA is shown.

The total length of the cDNA present in the recombinant clone studied is equal to 791 nucleotides. Only one of the 6 possible reading frames does not contain stop codons at a considerable length and the derived amino acid sequence (DAS) obtained in this reading frame, having 149 amino acids in the length, shows more than 90% homology with the amino acid sequence of αA_2-crystallin from a related frog species *Rana pipiens* [7]. The homology in this case was determined as the ratio of the number of identical amino acid residues at the corresponding positions to the total number of amino acid residues in each of the sequences compared. Since the amino acids 71–74 in the *R. pipiens* sequence are unknown, they were assumed to be identical in the sequences compared. This high homology provides a basis for the identification of pRt(1)297 as of a clone coding for αA_2-crystallin.

The conclusion that pRt(1)297 codes for αA_2-crystallin is further supported by its comparison with α-crystallins of the calf lens where the structure of both αA_2 and αB_2 polypeptides has been described [8]. This comparison demonstrates that the homology of DAS with calf lens αA_2 polypeptide is equal to 77% while its homology with αB_2 polypeptide amounts to 51%. In the latter case, it was necessary to introduce gaps in the αB_2 sequence, since the length of the αA_2 and αB_2 polypeptides differs by two amino acids. These gaps were counted as differences. Furthermore, the αA_2 and αB_2 crystallins markedly differ in their isoelectric points, which equal 5.9 for bovine αA_2 chain and 7.4 for bovine αB_2 chain. The pI value for the DAS determined using a computer program is equal to 5.9. This is in excellent agreement with the value for αA_2 crystallin, although it should be remembered that the DAS for pRt(1)297 cDNA lacks about 24 of αA_2 crystallin N-terminal amino acids, assuming that *R. pipiens* and *R. temporaria* αA_2-crystallins are identical in length. The region of mRNA corresponding to the first 24 amino acids and 5′-terminal non-translated region was apparently lost during the cloning procedure. In pRt(1)297 cDNA it is replaced by an artefactual sequence (positions from −1 to −185), which is complementary to the sequence 184–368 in the coding part. The appearance of the −1 to −185 region in cDNA is probably the result of erroneous copying during cDNA synthesis. Such observations were made earlier by us and by others [9–11].

The known structure of frog αA_2-crystallin cDNA and recent publications describing the structure of cloned rat and mouse αA_2 cDNA sequences

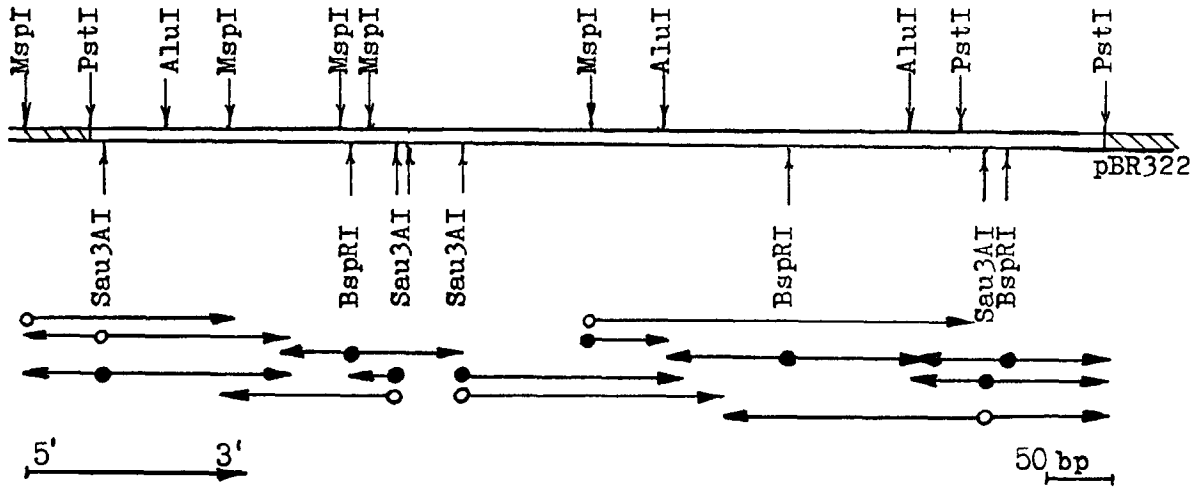

Fig.1. Cleavage map of αA_2-crystallin cDNA cloned in pRt(1)297 recombinant plasmid and the strategy of sequencing. The sequence was read from the corresponding sites labelled with [γ-^{32}P]ATP (●) or [α-^{32}P]dNTP (○); (⟶) direction of mRNA translation.

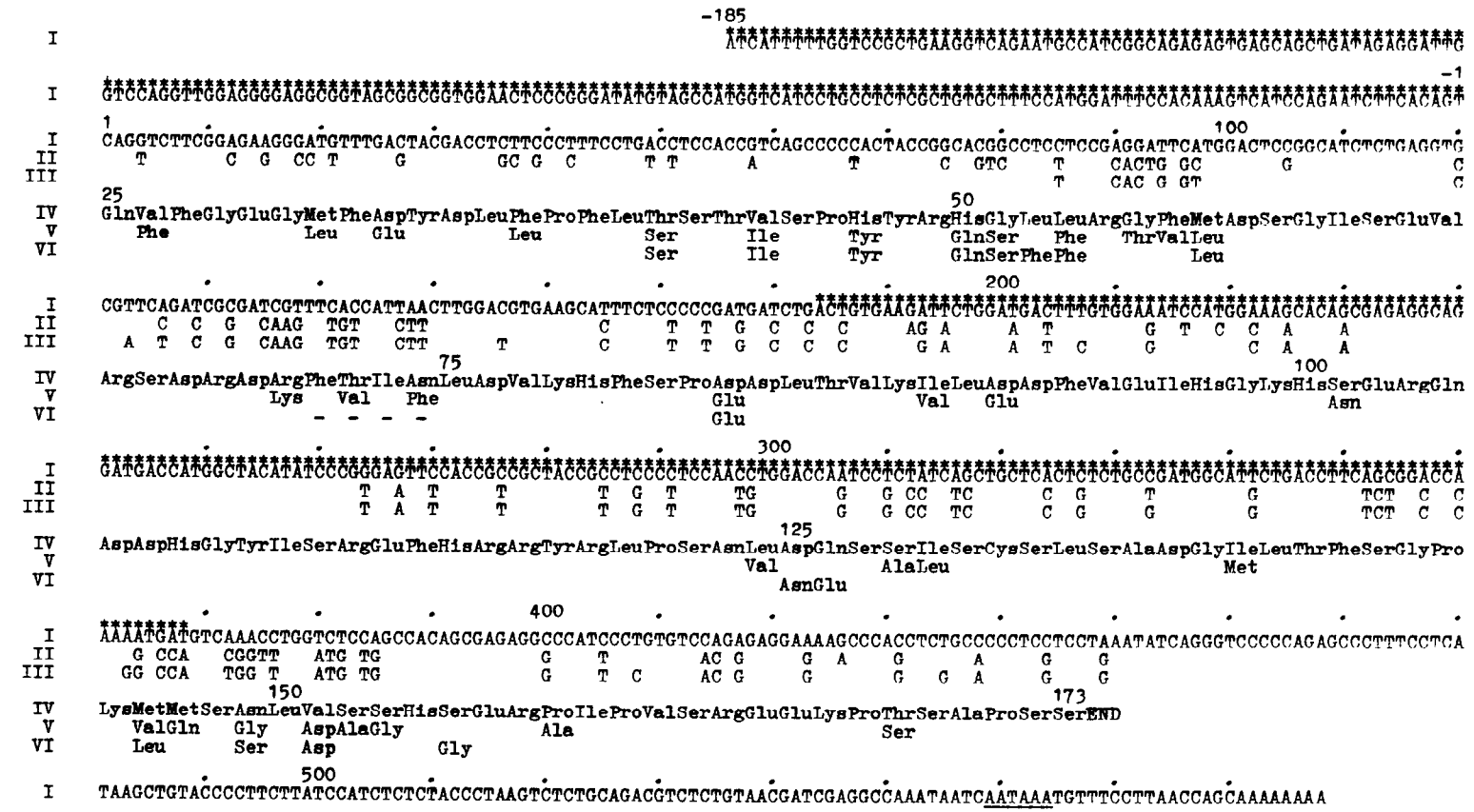

Fig.2. Nucleotide sequence of the frog (*Rana temporaria*) αA_2-crystallin cDNA (I) compared with the sequences of mouse (II) and rat (III) αA_2-crystallin mRNAs published in the literature (see text); (IV) derived amino acid sequence (DAS) of *R. temporaria* cDNA; (V) DAS of rodent αA_2-crystallin; (VI) published amino acid sequence of *R. pipiens* αA_2-crystallin. Note that the numbering of the αA_2-crystallin amino acid sequence is from the residue 25 since the region of frog mRNA corresponding to the first 24 amino acids has been lost during cloning. The artefactual sequence (positions from -1 to -185) at the 5′-end is marked with asterisks, as well as the complementary sequence 184–369 in the coding part. Universal polyadenylation sequence is underlined.

[12,13] allow the rate of their molecular evolution to be estimated. The coding region of the frog αA_2-cDNA (positions 1–447) shows a 75% homology with the nucleotide sequence of cloned murine αA_2-crystallin cDNA as well as with the nucleotide sequence of αA_2 rat crystallin (positions 83–447). In the latter case, the comparison was made from position 83 since the published structure of rat αA_2-crystallin cDNA does not contain the region corresponding to the first 52 amino acids. When comparing frog and mouse αA_2-crystallin sequences, it can be seen that 108 nucleotide substitutions result in changes in 32 amino acid residues in the molecule. The corresponding comparison of frog and rat αA_2 mRNA species reveals that 89 nucleotide substitutions result in changes in 23 amino acids. No extended homologies have been found in the 3′-NTR (see below).

The coding region of αA_2 crystallin does not contain internal duplications similar to those found in genes coding for γ- and β-crystallins [11,14,15]. The longest repeat is equal to 13 nucleotides (positions 226–238 and 388–400).

The main differences between the nucleotide sequences of frog and rodent αA_2-crystallin mRNAs are found in 3′-NTR (positions 451–580). The length of frog lens αA_2-crystallin 3′-NTR is equal to 130 nucleotides, that is less than 25% of mouse (536 nucleotides) and rat (583 nucleotides) 3′-NTR. The presence of a universal polyadenylation signal in frog lens 3′-NTR (positions 559–564) as well as of the poly(A) stretch, is an argument indicating that the 3′-NTR present in cloned cDNA of pRT(1)297 is intact.

The presence of a relatively short 3′-NTR in αA_2-crystallin mRNA of the frog implies that the length of this RNA should be markedly lower than that of the corresponding mRNA of mammals possessing a long 3′-NTR. Blot hybridization was performed (fig.3) to confirm this. Total frog lens poly(A)$^+$RNA was fractionated by electrophoresis, the gel was blotted onto nitrocellulose and the transfer hybridized with nick-translated [^{32}P]DNA of pRt(1)297 coding for αA_2-crystallin. The size of αA_2-crystallin frog mRNA was found to be equal to 700–800 nucleotides. In order to exclude the possibility that mRNA underwent fragmentation we hybridized a parallel lane from the same transfer with the labelled recombinant plasmid

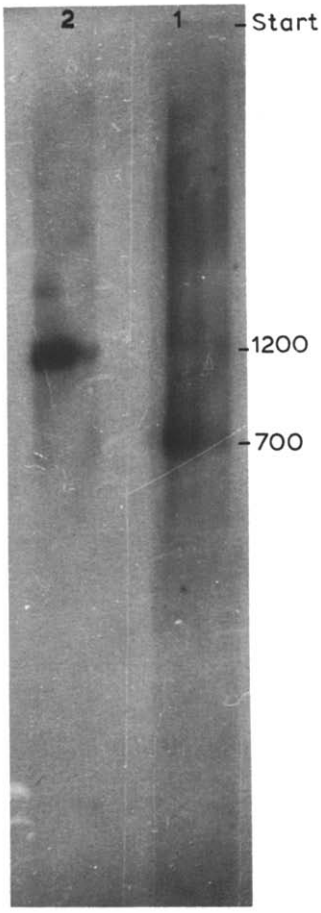

Fig.3. Blot hybridization of total frog lens poly(A)$^+$RNA with labeled pRt(1)297 plasmid coding for αA_2-crystallin (lane 1) and pRt(1)95 plasmid coding for 35 kDa β-crystallin (lane 2).

coding for the 35 kDa polypeptide from the frog lens. The length of the corresponding mRNA was found to be equal to 1200 nucleotides in agreement with its capacity to code for a 35 kDa polypeptide. We conclude that mRNA in this preparation of poly(A)$^+$RNA is intact and the estimated size of αA_2-crystallin mRNA reflects its true length. Thus frog lens αA_2-crystallin mRNA is almost twice as short as the corresponding mRNA of the other vertebrate species studied with a size equal to about 1400 nucleotides [12,13,16] and this difference appears to be due to the presence of the short 3′-NTR.

Two hypotheses may be suggested to explain the extreme difference in length between the 3′-NTR

of the frog lens and mammalian αA_2-crystallin mRNA:

(i) Assumes a different length of the corresponding genomic genes, possibly associated with a mutation in the transcription termination sequences during evolution from amphibia to mammals.

(ii) Assumes that the length of αA_2-crystallin transcription units in the frog and in mammals is not critically different, but the splicing site normally present in the 3′-non-coding region of the frog gene is lacking in the mammalian gene, leading to an exaggerated 3′-NTR of mammalian αA_2-crystallin mRNA.

These possibilities may be tested after the isolation of the frog genomic gene coding for αA_2-crystallin.

REFERENCES

[1] Bloemendal, H. (1982) CRC Crit. Rev. Biochem. 12, 1–38.
[2] Kozlov, K.A., Tomarev, S.I., Dzhumagaliev, E.B., Dolgilevich, S.M., Skobeleva, N.A., Frolova, L.Yu. and Gause, G.G. (1980) Dokl. Acad. Sci. USSR 255, 1503–1506.
[3] Tomarev, S.I., Dolgilevich, S.M., Kozlov, K.A., Zinovieve, R.D., Dzhumagaliev, E.B., Kogan, G.L., Skobeleva, N.A., Mikhailov, A.T., Frolova, L.Yu. and Gause, G.G. (1982) Gene 17, 131–138.
[4] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[5] Thomas, P. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[6] Maizel, J.V. and Lenk, R.P. (1981) Proc. Natl. Acad. Sci. USA 78, 7665–7669.
[7] De Jong, W.W., Gleaves, J.T. and Boulter, D. (1977) J. Mol. Evol. 10, 123–135.
[8] Van der Ouderae, F.J., De Jong, W.W., Hilderink, A. and Bloemendal, H. (1974) Eur. J. Biochem. 49, 157–168.
[9] Weaver, C.A., Gordon, D.F. and Kemper, B. (1981) Proc. Natl. Acad. Sci. USA 78, 4073–4077.
[10] Volckaert, G., Tavernier, J., Derynck, R., Devos, R. and Fiers, W. (1981) Gene 15, 215–219.
[11] Tomarev, S.I., Krayev, A.S., Skryabin, K.G., Bayev, A.A. and Gause, G.G. (1982) FEBS Lett. 146, 315–318.
[12] King, C.R., Shinohara, T. and Piatigorsky, J. (1982) Science 215, 985–987.
[13] Moormann, R.J.M., Van der Velden, H.M.W., Dodemont, H.J., Andreoli, P.M., Bloemendal, H. and Schoenmakers, J.G.G. (1981) Nucleic Acid. Res. 9, 4813–4822.
[14] Inana, G., Shinohara, T., Maizel, J.V. and Piatigorsky, J. (1982) J. Biol. Chem. 257, 9064–9071.
[15] Moormann, R.J.M., Den Dunnen, J.T., Bloemendal, H. and Schoenmakers, J.G.G. (1982) Proc. Natl. Acad. Sci. USA 79, 6876–6880.
[16] Bloemendal, H. (1977) Science 197, 127–138.