Vol. 190, No. 2, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 29, 1993 Pages 391-396

5'-RACE PCR of mRNA for Three Taxon-Specific Crystallins: For each gene one promoter controls both lens and non-lens expression

Jason Hodin and Graeme Wistow*

Section on Molecular Structure and Function,
6/222, LMDB, National Eye Institute,
National Institutes of Health, Bethesda, MD 20892

Received November 23, 1992

Summary: Gene recruitment of enzyme crystallins is a novel process in molecular evolution in which genes for some metabolic enzymes acquire extremely high expression in lens without prior gene duplication. Using the RACE method, the 5' end of the mRNAs for duck lens δ1-, argininosuccinate lyase/δ2- and lactate dehydrogenase-B/ε-crystallin has been amplified from lens and non-lens tissues and sequenced. In all three cases the major transcription start sites were identical in lens and in other tissues. This suggests that for these three genes, and in contrast to at least one other example, gene recruitment does not require the presence of alternative tissue-preferred promoters.

The highly abundant crystallins of the vertebrate eye lens have arisen by recruitment of pre-existing proteins through modification of gene expression (1-3). This phenomenon of gene recruitment requires no prior gene duplication with the result is that one protein acquires two distinct functions. This is in contrast to the classical model in which gene duplication must occur before new protein functions arise (4). Many recruited crystallins, whether ubiquitous stress crystallins, such as the small heat shock protein αB-crystallin, or taxon-specific enzyme crystallins such as lactate dehydrogenase-B (EC 1.1.1.27)(LDH)/ε-crystallin, have now been characterized (2,3,5). However, no uniform mechanism for lens preferred expression has yet emerged.

An important recent discovery has revealed that the gene for one taxon-specific crystallin, orthoquinone reductase/ζ-crystallin of the guinea pig (6), has alternative promoters, one of which confers abundant, lens-preferred expression (7,8). The possibility that this represents a general mechanism is hinted at by other features of known enzyme crystallin genes, including the presence of additional 5' UTR (untranslated region) exons in

*To whom correspondence should be addressed.

0006-291X/93 $4.00

δ-crystallin genes when compared to non-recruited mammalian homologues, argininosuccinate lyase (EC 4.3.2.1) (ASL)(9-12). Since alternative promoters reveal their existence by conferring differing 5' terminal regions on gene transcripts we have examined the 5' ends of the mRNA of three avian taxon-specific crystallin genes for evidence of tissue-preferred alternative promoters.

Materials and Methods

RNA preparation and cDNA synthesis- Total RNA was extracted from 14-day embryo duck (*Anas platyrhynchos*) tissues (Truslow Farms, Chestertown, MD) by the guanidinium thiocyanate method (13) using RNAsol (Cinna/Biotecx, Friendswood, TX). First strand cDNA templates were synthesized using Copy Kit (Invitrogen, San Diego CA) following manufacturers instructions. Various strategies were adopted for priming templates. 5μg of total RNA were used in all cases. Duck lens and heart templates were synthesized by random priming. Brain and liver templates were produced with a primer specific to δ-crystallin but common to both mRNAs (14) (6426: AGGGAATTCCTCAGAGATCTTTTCCAGGCCA). Final volume of template in all cases was 200μl.

RACE- For rapid amplification of cDNA ends (RACE), the methodology of Frohman (15) was employed. Templates were tailed with poly(A) using terminal deoxytransferase (TdT) (BRL, Gaithersburg, MD). Sequences corresponding to the 5' ends of mRNA molecules were then amplified by the polymerase chain reaction (PCR) using specific primers at one end and oligo(dT) (primer 6672: AGGAATTCGCTACTCGAG$(T)_{30}$) at the other. Primers were synthesized using an Applied Biosystems 380B DNA synthesizer. All primers were designed with restriction sites to facilitate subsequent cloning. Typical PCR reactions used 5μl of template, 0.4μM of each primer, 0.5mM of each dNTP, 1.25mM $MgCl_2$, 2.5U AmpliTaq DNA polymerase (Perkin Elmer Cetus, Norwalk, CT), in 100μl total volume. PCR program was 94C 2 min; 30x[94C 1.5min; 55C 2min; 72C 3min]; 72C 10min; 4C. ε-crystallin was primed with primer 6425 (GAGGAATTCCGCCTTTGAATACACCCACATTCCT), taken from the published cDNA sequence (16). For lens, δ1 and δ2 were primed with the common primer 6426 (AGGGAATTCCTCAGAGATCTTTTCCAGGCCA). Specificity was achieved at the sequencing step. For brain both δ cDNAs were primed with primer 6471 (CTCGAGATCTTCTCCAGTTCAGT). For liver the same procedure was applied then, in an attempt to improve results, specific δ1 (6969: CATGATGGGATCCGTGCTTCCGAC) and δ2 (6970: ATGATGGGATCTGTGCTTCCACTG) primers were used.

Sequencing- PCR fragments were isolated by gel purification from low melting temperature agarose using MagicPrep kit (Promega, Madison, WI). Purified fragments were sequenced in a thermocycler using Taq polymerase and other materials and protocols of the *fmol* DNA sequencing system (Promega). Improved results were obtained using specific primers at lower concentrations than recommended, 25ng per reaction. For ε-crystallin from lens and heart primer 6115 (CTGCCCGACGCCCACCACCGTTAT) was used. For δ1 from lens and brain primer 6969 was used (as above). For δ2 from lens and brain primer 6970 was used (as above). Neither preparation of liver PCR fragments produced readable sequence by this method for δ1 or δ2.

To obtain liver sequence and to confirm other δ1, δ2 and ε sequences, fragments were subcloned into pBluescript (Stratagene, La Jolla, CA), making use of Eco RI, Xho I and other restriction sites included in the primers. Cloning followed standard methodologies (17). Sequencing used the Sequenase kit and methods (USB, Cleveland OH). New sequences

for duck δ-crystallins are to be included in existing GenBank entries with accession numbers M35133 and M35132. New sequence for duck LDH-B/ε-crystallin has the accession number Z15016.

Results and Discussion

The acquisition of alternative promoters is attractive as a general mechanism for gene recruitment. It would allow uninterupted function of the existing enzyme promoter while the novel lens promoter was evolving, thus minimizing the risk of loss or disruption of normal gene expression which might otherwise occur during modification of a single promoter.

Indeed, one example of the use of alternative promoters has recently been discovered. The gene for guinea pig ζ-crystallin has two separate promoters (7,8), one of which appears to be responsible for non-lens expression while the other, a TATA-box promoter, is strongly lens-preferred both in cultured cells and in transgenic mice (8). In contrast, there is a rather specialized case in which a single promoter seems to have both lens and non-lens expression. An 800bp fragment of the duck α-enolase/τ-crystallin promoter is highly active but exhibits no lens-preference either in cultured cells (18) or in adult transgenic mice (19). This particular enzyme crystallin is somewhat atypical, having preferred expression in lens epithelia rather than fiber cells and a narrow window of enhanced expression during development which may be difficult to mimic in model systems (18-20).

Avian δ2-crystallins are active argininosuccinate lyases (ASL) (21,22) and may also be major crystallins (14). Presumably as a result of the pressures of adaptive conflict (22, 23) the ancestral ASL/δ-crystallin gene in the bird lineage duplicated to generate the nonenzymatic δ1-crystallin, specialized for lens. In the chicken, both δ-crystallin genes have TATA-box promoters and 17 exons (9,10) while mammalian ASL genes have 16 exons and lack TATA-boxes (11,12). Furthermore, the first intron of the chicken δ2-crystallin/ASL gene contains a sequence, (CAGGGAAGG), similar to a consensus (CAGGGCAGGCG/TGCTC) common to the promoters of mammalian urea cycle genes (12). This suggests that the recruitment of δ-crystallin involved the acquisition of a novel TATA-box promoter upstream of an existing promoter orthologous to that of mammalian ASL genes and, consequently, that a non-lens promoter could lie downstream in the first intron of δ-crystallin genes. Lens-preferred expression of chicken δ1-crystallin has been attributed to an enhancer in intron 3 rather than to the 5' promoter region (24). However, the possibility of alternative promoters which could be responsible for expression of δ-crystallins in non-lens tissue does not appear to have been addressed directly.

The possibility that δ-crystallins may have been recruited through the use of alternative promoters was investigated by analyzing the 5' UTR sequences of δ1- and δ2-crystallins in duck lens, brain and liver. The RACE PCR method (15) was used to visualize

transcripts in very low abundance in non-lens tissues and to allow sequencing. In embryonic chicken there are detectable, but extremely low, levels of δ1- and δ1-crystallin mRNA in non-lens tissues (25). The same procedure was applied to LDH-B/ε-crystallin in duck lens and heart, since no 5' UTR sequence for this enzyme crystallin had been published. The duck was chosen for this purpose, since it expresses very high levels of LDH-B/ε-crystallin (1) and both δ1- and ASL/δ2-crystallins (14) in lens.

First strand CDNA for these crystallins was synthesized from mRNA using either random priming or crystallin-specific synthetic primers. This is essentially equivalent to the conventional process of 5'-end mapping by primer extension (17). The cDNA was tailed with poly(A) and amplified using oligo-dT and specific primers. Template free controls were blank and independently derived templates gave similar results, suggesting that there was no cross-contamination of PCR products between experiments. Direct sequencing of PCR products was used to visualize and identify 5' termini in lens, brain and heart. Major sequences were unambiguous up to the synthetic 5' poly(A) tails. The attachment site for the synthetic tail often exhibited one or two bases of "wobble", overlapping with the cDNA sequence. Presumably this resulted from slightly premature termination of the reverse transcription reaction or some variation in the initiation of gene transcription itself. To confirm the direct sequences, PCR fragments were subsequently subcloned into a plasmid vector and resequenced. The error rate from PCR was very low in these short sequences (close to zero).

No sequence has previously been published for the 5' untranslated region or initiator methionine of LDH-B/ε-crystallin mRNA. RACE PCR showed identical 5' sequences for ε-crystallin from duck lens and heart (fig. 1a). There was no evidence of significant alternative start sites in either tissue. The 5' untranslated region was relatively short, only 40nt in length.

For duck δ-crystallins some 5' untranslated sequence was already known from cDNA cloning (14). This sequence was confirmed by RACE PCR which again showed identical 5'

a) 5' ACGGCTCTGCATCCCCCTGCTCACAGCACTAACGGACGCCATGgcg...

b) 5' ATTTACGGAGCTACCGGCCAgggctccactgcagacacg...

c) 5' ACACGGAGCCAccggggctgcactgcggacacg.....

Fig. 1. 5' sequences of enzyme crystallins from lens and non-lens tissues derived from direct sequence and from cloned fragments. Lower case indicates start of published cDNA sequence.
a) 5' sequence of duck LDH-B/ε-crystallin from lens and heart.
b) 5' sequence of duck δ1-crystallin from lens and brain.
c) 5' sequence of duck δ2-crystallin from lens and brain.

ends for δ1 in lens and brain and for δ2 in lens and brain (figs. 1b,c). The unusual six-base insertion in δ2-crystallin, an example of splice junction slippage (14), was confirmed. Since the very different transcripts for mammalian ASL genes were identified in liver, δ-crystallins were also examined in duck liver. Identically sized products to those from lens were obtained, but for unknown reasons did not yield readable direct sequence. Fragments were therefore subcloned for sequencing. Cloned liver transcripts were more variable than those from other tissues, perhaps reflecting a problem with RNA quality. Even so, full-length sequence for δ1-crystallin, identical to that from lens, was observed. There were also two full-length sequences for δ2-crystallin from liver although additional shorter clones were observed. Two initiated in exon 1, upstream of any putative intron 1 promoter and one began in exon 2. Thus the possibility of some transcription from a downstream promoter could not be ruled out, although the 5' UTR exon from lens is clearly also present in liver. Shorter fragments are most likely the result of mRNA truncation.

Thus, in spite of the intriguing features of δ-crystallin gene structure and the presence of two promoters in ζ-crystallin promoters, the results described here suggest that tissue-preferred alternative promoters are not a general feature of the enzyme crystallin recruitment process. For three different taxon-specific crystallin genes, δ1-, δ2- and ε-crystallins, the same promoter which confers high expression in lens also functions in other embryonic tissues. In these cases no other promoters can be seen to be making a large contribution to non-lens gene expression for any of these genes in embryonic duck tissues. While δ-crystallin recruitment may initially have involved the acquisition of a novel upstream promoter there is no evidence that the putative ancestral enzyme promoter still functions. Whatever their history, δ-crystallin genes do not now seem to have alternative promoters.

Similar results were obtained for LDH-B/ε-crystallin in embryonic duck, in which the same promoter responsible for high expression in the lens also produces transcripts in the heart. When these results were obtained there was no information on gene structure or 5' UTR sequence for this enzyme crystallin. However, the same 5' UTR for duck lens ε-crystallin has recently been described by others in a meeting abstract (26). These workers found the same initiation site being used in lens and heart but, by primer extension, also found an additional start site 28bp further upstream in adult heart. Our results gave no evidence for such a additional start site in embryonic heart. This may reflect a developmental phenomenon in heart, but in any case does not alter the conclusion that the lens promoter is also used in other tissues. Hence, use of alternative promoters *per se* is not responsible for gene recruitment in this case.

Instead, it seems likely that for τ-, ε-, δ1- and δ2-crystallins high expression in lens depends on a single promoter, perhaps modified with a TATA box, lens-specific *cis*-

elements, or lens-specific enhancers elsewhere in the gene to give rise to a new mode of expression in lens. Indeed, the diversity of crystallins themselves may be reflected in a diversity of mechanisms for their recruitment.

References

1. Wistow,G.J., Mulders,J.W. and de Jong,W.W. (1987) *Nature*, **326**, 622-62
2. Wistow,G. and Piatigorsky,J. (1988) *Ann. Rev. Biochem.*, **57**, 479-504.
3. Piatigorsky,J. and Wistow,G. (1991) *Science*, **252**, 1078-1079.
4. Kimura,M. and Ohta,T. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 2848-2852.
5. de Jong,W.W., Hendriks,W., Mulders,J.W. and Bloemendal,H. (1989) *Trends Biochem. Sci.*, **14**, 365-368.
6. Rao,P.V., Krishna,C.M. and Zigler,J.S. (1992) *J. Biol. Chem.*, **267**, 96-102.
7. Gonzalez,P., Rao,P.V., Hernandez-Calzadilla,C., Borras,T. and Zigler,J.S. (1992) *Exp. Eye Res.*, **55**(Sep Suppl. 1), 199-199.
8. Lee,D.C., Gonzalez,P., Rao,P.V., Zigler,J.S.,Jr. and Wistow,G.J. (1992) *Enzymology and Molecular Biology of Carbonyl Metabolism*, in press.
9. Nickerson,J.M., Wawrousek,E.F., Hawkins,J.W., Wakil,A.S., Wistow,G.J., Thomas,G., Norman,B.L. and Piatigorsky,J. (1985) *J. Biol. Chem.*, **260**, 9100-9105.
10. Nickerson,J.M., Wawrousek,E.F., Borras,T., Hawkins,J.W., Norman,B.L., Filpula,D.R., Nagle,J.W., Ally,A.H. and Piatigorsky,J. (1986) *J. Biol. Chem.*, **261**, 552-557.
11. Matsubasa,T., Takiguchi,M., Amaya,Y., Matsuda,I. and Mori,M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 592-596.
12. Abramson,R.D., Barbosa,P., Kalumuk,K. and O'Brien,W.E. (1991) *Genomics*, **10**, 126-132.
13. Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156-159.
14. Wistow,G. and Piatigorsky,J. (1990) *Gene*, **96**, 263-270.
15. Frohman,M.A. (1990) In *PCR Protocols: A Guide to Methods and Applications*(Innis,M.A., Gelfand,D.H., Sninsky,J.J. and White,T.J. eds.), . Academic Press, San Diego, CA, pp. 28-38.
16. Hendriks,W., Mulders,J.W.M., Bibby,M.A., Slingsby,C., Bloemendal,H. and de Jong,W.W. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7114-7118.
17. Davis,L.G., Dibner,M.D. and Battey,J.F. (eds.) (1986) *Basic Methods in Molecular Biology*. Elsevier, New York, NY.
18. Kim,R.Y., Lietman,T., Piatigorsky,J. and Wistow,G.J. (1991) *Gene*, **103**, 193-200.
19. Kim,R.Y. and Wistow,G.J. (1992) , submitted.
20. Wistow,G.J., Lietman,T., Williams,L.A., Stapel,S.O., de Jong,W.W., Horwitz,J. and Piatigorsky,J. (1988) *J. Cell Biol.*, **107**, 2729-2736.
21. Kondoh,H., Araki,I., Yasuda,K., Matsubasa,T. and Mori,M. (1991) *Gene*, **99**, 267-271.
22. Barbosa,P., Wistow,G.J., Cialkowski,M., Piatigorsky,J. and O'Brien,W.E. (1991) *J. Biol. Chem.*, **266**, 22319-22322.
23. Geliebter,J. (1987) *BRL Focus*, **9**, 5-8.
24. Hayashi,S., Goto,K., Okada,T.S. and Kondoh,H. (1987) *Genes Dev.*, **1**, 818-828.
25. Thomas,G., Zelenka,P.S., Cuthbertson,R.A., Norman,B.L. and Piatigorsky,J. (1990) *New Biologist*, **2**, 903-914.
26. Schoenmakers,J.G.G., Kraft,H.J., Hendriks,W., de Jong,W.W. and Lubsen,N.H. (1992) *Exp. Eye Res.*, **55**(Sep Suppl. 1), 222-222.