Gene, 167 (1995) 321–325

GENE 09342

Isolation and characterization of the 5′-flanking sequence of the human ocular lens *MIP* gene

(Eye lens fiber membrane; major intrinsic protein; promoter; transcriptional regulation; regulatory element; *Alu* repeat)

Xiao Yan Wang, Chiaki Ohtaka-Maruyama*, M. Michele Pisano*, Cynthia J. Jaworski and Ana B. Chepelinsky

Laboratory of Molecular and Developmental Biology, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received by Z.E. Zehrer: 3 April 1995; Revised/Accepted: 21 July/2 August 1995; Received at publishers: 11 September 1995

SUMMARY

The *MIP* (major intrinsic protein) gene, a member of an ancient family of membrane channel genes, encodes the predominant fiber cell membrane protein of the ocular lens. Its specific expression in the lens fibers is temporally and spatially regulated during development. To study the regulation of expression of *MIP* and delineate the regulatory elements underlying its tissue specificity and ontogenic profile, we have cloned 2840 bp of the human *MIP* 5′-flanking sequence. The human *MIP* 5′-flanking sequence contains three complete *Alu* repetitive elements in tandem at position between nt −1699 and −2684 (nt −1699/−2684). These *Alu* elements appear to have had a complex evolutionary history with insertions at different times. We have fused DNA fragments containing *MIP* 5′-flanking sequences to the bacterial *cat* reporter gene encoding chloramphenicol acetyltransferase and assayed them in primary cultures of chicken lens cells. We have mapped two negative regulatory regions in the human *MIP* 5′-flanking sequences −1564/−1696 and −948/−1000. We demonstrated that the human *MIP* 5′-flanking sequence −253/+42 contains a functional promoter in lens cells but is inactive in kidney epithelial cells or mouse fibroblasts, suggesting that this sequence contains regulatory elements responsible for the lens-specific expression of *MIP*.

INTRODUCTION

The *MIP* gene is specifically expressed in the ocular lens and encodes the most abundant protein of the lens fiber membrane. *MIP* expression exhibits temporal and spatial regulation during lens development. Its expression is first detectable in the primary lens fibers at the early lens vesicle stage and subsequently in secondary lens fibers as they differentiate from epithelial cells (Yancey et al., 1988). *MIP* belongs to an ancient family of transmembrane channel proteins that arose by gene duplication of an ancestral gene. Purified MIP forms large channels when incorporated into liposomes and voltage-dependent channels in planar lipid bilayers. During cataractogenesis and aging, MIP undergoes selective proteolysis which may modify the gating of the MIP channels

Correspondence to: Dr. A.B. Chepelinsky, National Institutes of Health, Building 6, Room 211, 6 Center DR MSC 2730, Bethesda, MD 20892-2730, USA. Tel. (1-301) 496-9615; Fax (1-301) 402-0781; e-mail: abc@helix.nih.gov

* Present addresses: (C.O.-M.) Cellular Physiology Laboratory, The Institute of Physical and Chemical Research, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-01, Japan. Tel. (81-48) 462-1111; (M.M.P.) Department of Pathology, Anatomy and Cell Biology, Jefferson Medical College, Thomas Jefferson University, 1020 Locust Street, Philadelphia, PA 19107, USA. Tel. (1-215) 955-7855.

Abbreviations: bp, base pair(s); βGal, β-galactosidase; CAT, chloramphenicol acetyltransferase; *cat*, gene encoding CAT; *cry*, gene(s) encoding crystallin(s); CPLE, chicken primary lens epithelium; kb, kilobase(s) or 1000 bp; MIP, major intrinsic protein of the ocular lens; *MIP*, gene encoding MIP; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; *SINE*, short interspersed element(s); *tsp*, transcription start point(s).

SSDI 0378-1119(95)00637-0

(see, for review, Ehring et al., 1993; Chepelinsky, 1994).

The ability of the lens to refract and focus incoming light onto the retina depends upon the large concentration of soluble structural proteins called crystallins, as well as the unique cytoarchitecture of the mature fiber cells. The intercellular space between fibers is extremely narrow, an essential property for maintaining the transparency and correct refractive index of the lens (see Duncan and Jacob, 1984; Harding and Crabbe, 1984). The spatial localization of MIP within the ocular lens and its properties as a channel suggest its possible role in the maintenance of lens transparency by reducing the interfiber space (Ehring et al., 1993). While the regulation of *cry* expression in the lens has been extensively studied (see, for review, Piatigorsky and Zelenka, 1992), little is known about the regulation of expression of genes encoding lens membrane proteins. Thus, we have been investigating the expression of *MIP* in order to determine whether any common features are shared with *cry* expression in the lens. We previously isolated a clone containing the human *MIP* and characterized the structural gene (Pisano and Chepelinsky, 1991). Here we report the characterization in the human *MIP* 5′-flanking sequence of a functional promoter in lens cells.

EXPERIMENTAL AND DISCUSSION

(a) Isolation and characterization of a clone containing the 5′-flanking sequence of the human *MIP*

A genomic clone containing the human *MIP* and its 5′-flanking sequence was previously isolated and the sequence and characterization of the structural gene has been reported (Pisano and Chepelinsky, 1991). The nt sequence of the 5′-flanking region of the human *MIP* has now been determined by DNA sequencing of the pHMIP1, pHMIP2 and pHMIP3 clones, generated from the original *MIP* genomic clone, pHMIP16 (Pisano and Chepelinsky, 1991). 2840 bp of human *MIP* 5′-flanking sequence are shown in Fig. 1.

Primer extension of human lens RNA mapped the *tsp* of the human *MIP* 26 bp downstream from the TATA box (Pisano and Chepelinsky, 1991). The *tsp* allows us to define the non-transcribed 5′-flanking sequences that may contain *cis* regulatory elements of the human *MIP*.

(b) *Alu* elements in the human *MIP* 5′-flanking region

Computer analysis of the human *MIP* 5′-flanking sequence indicates the presence of four tandemly arranged *Alu* repetitive elements between positions −1699 and −2840. Three of them are full-length and the fourth one, located at the 5′ end of the isolated human *MIP* genomic clone (position −2791/−2840) is incomplete; only 51 bp are present.

The individual characteristics of the three complete *Alu* elements associated with the human *MIP* 5′-flanking sequence are listed in Table I. The *Alu* elements at nt −2684/−2408 and −2340/−2036 are in the sense orientation. The proximal *Alu* element at nt −1994/−1699 and the partial element at nt −2791/−2840 have antisense orientations. The three full-length elements contain poly(A) tails and are flanked by perfect direct repeats, in keeping with the retrotransposon origin of this family of inserted sequences (Rogers, 1985). According to the classification scheme of Jurka and Milosavljevic (1991), the element at −2340/−2036 belongs to the most ancient group of primate *Alu* repeats, the J family. The *Alu* at nt −2684/−2408 belongs to the Sx subgroup of the S family and has a 15-bp deletion at position −2489/−2490 (position 197/211 of the consensus *Alu* sequence) and a 2-bp deletion at position −2470/−2471 (231/232 of the consensus *Alu* sequence). The *Alu* element at nt

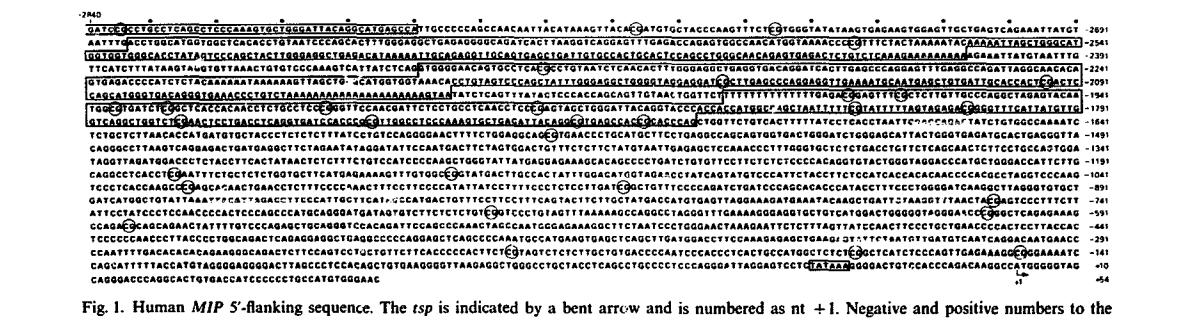

Fig. 1. Human *MIP* 5′-flanking sequence. The *tsp* is indicated by a bent arrow and is numbered as nt +1. Negative and positive numbers to the right indicate nt position upstream and downstream from the *tsp*, respectively. The *Alu* repeats and the TATA box are boxed. CG sequences are circled. This sequence has been entered into the GenBank database under the accession No U36308. **Methods:** Human *MIP* 5′-flanking sequence was established by sequencing subcloned fragments from a 16-kb *MIP* genomic clone, previously isolated from a human leukocyte genomic library, in M13 with universal primers as described in Pisano and Chepelinsky (1991). Sequence analyses were performed using Genetics Computer Group (Madison, WI, USA) software programs on the Frederick Cancer Research Facility VAX computer.

TABLE I

Alu elements of the human *MIP* gene 5′ flanking region

Position (from nt to nt)	Orientation[a]	Subfamily[b]	Identity[b] (%)	Flanking repeats (5′ to 3′)	CpG[c]
−2684/−2408	+	Sx	87	AGAAATTATGTAATTTG	1
−2340/−2036	+	J	79	TATCTCAG	3
−1994/−1699	−	Spq	90	CTGGTTCT	12

[a] + and − indicate that the repetitive element is located on either the coding or noncoding strand of the gene, respectively.
[b] The *Alu* repetitive elements within the human *MIP* locus are compared to the *Alu* consensus sequence and classified according to the scheme of Jurka and Milosavljevic (1991).
[c] Number of CpG dinucleotides per *Alu*.

−1994/−1699 shares characteristics of both the Sp and Sq families, and contains single bp insertions at positions 128/129 and 266/267 of the *Alu* consensus sequence, corresponding to positions −1825 and −1963. Thus, the cluster of *Alu* elements apparently arose by sequential insertions of distinct *Alu* source genes (Rogers, 1985). The J-family element at nt −2340/−2036, which shows the most divergence from the consensus (79% identity), appears to be the earliest insertion. The element at nt −1994/−1699, with the highest identity to the consensus (90%) and highest CpG content (12 CpG dinucleotides), is the most recent *Alu* element in this array.

Alu repetitive elements may modify transcription of nearby genes and repress transcription when methylated (Howard and Sakamoto, 1990; Englander et al., 1993). CpG methylation density has been shown to be responsible for transcriptional repression (Meehan et al., 1992; Hsieh, 1994) and methylation of γ-*cry* genes, a lens-specific gene family, may repress promoter activity (Peek et al., 1991). Therefore, we analyzed the possible role of the *Alu* repeats present in the 5′-flanking region of the human *MIP* in its transcriptional regulation.

(c) Negative and positive regulatory domains in the 5′-flanking region of the human *MIP*

Transient assays using the *cat* reporter gene were used to assess the transcriptional activity of the 5′-flanking sequences of the human *MIP*. A DNA fragment containing 2840 bp of human *MIP* 5′-flanking sequence and 42 bp of exon 1 was fused to *cat* (plasmid pHMIP2840-CAT). When CPLE cultures were transfected with this construction, no significant activation of *cat* expression was observed (see Fig. 2). Therefore, a shorter DNA fragment containing human *MIP* 5′-flanking sequence was also tested for promoter activity. A DNA fragment containing 253 bp of 5′-flanking sequence and 42 bp of exon one of the human *MIP* was fused to *cat*. The resulting construction (plasmid pHMIP253-CAT), activates *cat* expression when transfected into CPLE cultures (see Fig. 2), indicating the

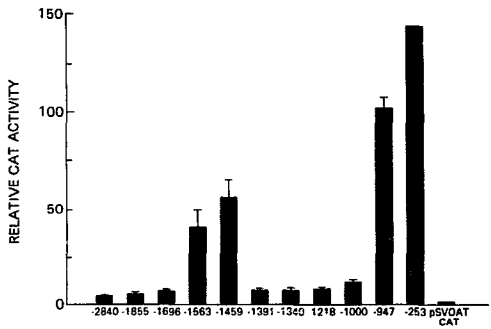

Fig. 2. Deletion analysis of human *MIP* promoter activity in transfected chicken lens cells primary cultures (CPLE). Bars indicate CAT activity relative to pSVOATCAT, for each deletion normalized to βGal activity. The length of human *MIP* 5′-flanking sequence fused to *cat* is indicated under each bar. The results plotted are the mean of at least triplicate transfections in two or more separate primary culture experiments, and the corresponding standard deviation is shown. **Methods:** Plasmids pHMIP253-CAT and pHMIP2840-CAT contain human MIP sequences at nt −253/+42 and −2840/+42, respectively, isolated from pHMIP3 (Pisano and Chepelinsky, 1991) and inserted upstream from *cat* in the pSVOATCAT expression vector (Meakin et al., 1989). Deletions at nt −1855, −1563, −1459 and −947 were constructed by deleting sequences −2840/−1856, −2840/−1564, −2840/−1460 and −2840/−948, respectively, from pHMIP2840-CAT. Deletion −1696 was constructed by ligating the PCR product containing human *MIP* sequence −1696/−1564 to the 5′ end of sequence −1563/+42 of pHMIP1563-CAT. Deletions −1391, −1340 and −1218 were constructed by ligating the PCR products containing human *MIP* sequence −1391/−948, −1340/−948 or −1218/−948 to the 5′ end of sequence −947/+42 of pHMIP947-CAT, respectively. Deletion −1000 was constructed by ligating a synthetic double-stranded oligo corresponding to sequence −1000/−948 to the 5′ end of sequence −947/+42 of pHMIP947-CAT. The sequences of the resulting constructions, containing the *MIP* sequences in the same orientation as in the endogenous gene, were confirmed by DNA sequencing. CPLE cultures and transfections were performed as previously reported (Borras et al., 1988). CPLE cultures were cotransfected with 10 μg *cat* expression vectors and 1.5 μg pSV-βGal vector (from Promega) for normalization of results. CAT enzymatic assay was performed by the biphase method (Neuman et al., 1987). βGal enzymatic assay was performed as described by Nielsen et al. (1983). CAT enzymatic activity was normalized to βGal activity and expressed relative to the activity of pSVOATCAT.

presence of a functional promoter in the human *MIP* sequence −253/+42. As these results suggest that the human *MIP* sequence −2840/−254 may contain negative regulatory elements of the human *MIP* promoter, further deletions in the 5′-flanking sequence of the human *MIP* were fused to *cat* and their ability to activate *cat* expression was assayed in transient assays in lens cells. As shown in Fig. 2, when −1855 or −1696 bp of 5′-flanking sequence are present, a very inefficient promoter activity is detected. In contrast, when −1563 or −1459 bp are fused to *cat* the efficiency of the promoter increases approx. 10-times. However, when only 1391 bp of 5′-flanking sequence are present, the efficiency of the promoter is drastically decreased again. Similar activities are observed with 1340, 1218 or 1000 bp of 5′-flanking sequence. When only 947 bp of 5′-flanking sequence are fused to *cat* the promoter activity increases approx. 10-times and functions comparably to the 253 bp of 5′-flanking sequence (see Fig. 2). These results suggest the presence of two negative regulatory domains in the *MIP* 5′-flanking sequence. One of them comprises the sequence −1696/−1564; the 5′ end of this domain may or may not include the *Alu* repeats. The other comprises the sequence −1000/−948; the precise delineation of the 5′ border of this domain in the region −1458/−1000 requires further studies. These two negative regulatory domains may modulate the positive regulatory elements present in the −1563/−1459 and −253/+42 regions of the human *MIP*.

(d) Regulatory sequence motifs present in the human *MIP* promoter

The pHMIP253-CAT plasmid activates *cat* expression when transfected into lens cells primary cultures (CPLE) but is inactive when transfected into RK-13 (rabbit kidney epithelial cells) or NIH-3T3 mouse fibroblasts (see Fig. 3). Therefore, the *MIP* sequence −253/+42 contains regulatory elements that direct *cat* expression to lens cells but are not functional in transfected non-lens cells. As numerous regulatory elements have been shown to be involved in *cry* gene expression in the lens, we examined the human *MIP* sequence −253/+42 to determine whether any common features are shared between *cry* and non-*cry* gene expression in the lens.

The sequence 5′-GGGAAATCCC of the mouse αA-*cry* promoter has been characterized as the binding site for αA-CRYBP1, a transcription factor related to the NF-κB/Rel family (Nakamura et al., 1990). Two putative binding sites for NF-κB/Rel (Baeuerle, 1991) are located in the human *MIP* sequence containing a functional promoter in lens cells; one at nt positions −15/−24, 3′ to the TATA box (5′-GGGACTGTCC), and the other at nt position −140/−149 (5′-GGGAAAATCC). Whether

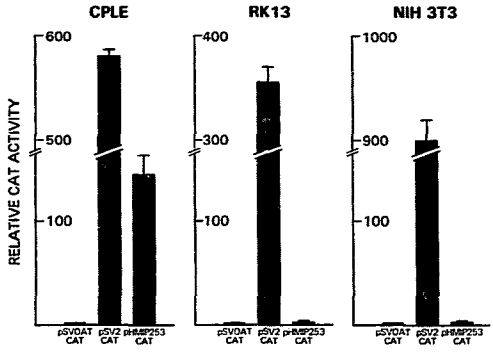

Fig. 3. Human *MIP* promoter activity in transfected lens cells and non-lens cells. Bars indicate pHMIP253-CAT or pSV2-CAT relative CAT activity to pSVOATCAT normalized to βGal activity cotransfected in CPLE cells (left panel), RK13 cells (middle panel) or NIH-3T3 cells (right panel). The results plotted are the mean of at least triplicate transfections and the corresponding standard deviation is indicated. 10 μg plasmids pHMIP253-CAT, pSV2-CAT (Gorman et al., 1982) or pSVOATCAT were cotransfected with 1.5 μg pSV-βGal for normalization of results and CAT enzymatic assay was performed as described in the legend to Fig. 2.

αA-CRYBP1 or another member of the NF-κB/rel transcription factor gene family is involved in regulating *MIP* transcription in the lens requires further study.

The mouse γF-*cry* sequence −48/−35, identified as the γF-1 binding motif and as the γ-box motif in the rat γD-*cry*, interacts with lens nuclear factors (Liu et al., 1991; Peek et al., 1992). This motif is part of the 43 bp region conserved in all members of the lens fiber-specific γ-*cry* gene family (Lok et al., 1989; Peek et al., 1990). The sequence motifs 5′-CTGCTACCTCAG and 5′-CTGTCCACCCAG, located at nt −61/−72 and −9/−20, respectively, of the human *MIP* have similarities to the γ-box consensus sequence, suggesting the possible involvement of a transcription factor that regulates expression of both *MIP* and γ-*cry* genes in the lens.

The presence of positive and negative regulatory elements has also been found in other genes expressed in lens cells: γD-*cry* (Peek et al., 1992), γF-*cry* (Lok et al., 1989), αA-*cry* (Sax et al., 1994) and δ1-*cry* (Borras et al., 1988; Kamachi and Kondoh, 1993). Further characterization of the cis regulatory elements of the human *MIP* will allow us to understand how their interaction with general and lens-specific transcription factors regulates the temporal and spatial expression of *MIP* in the ocular lens.

(e) Conclusions

(*1*) We have cloned 2840 bp of the human *MIP* 5′-flanking sequence and found that it contains three

complete *Alu* repetitive elements in tandem, located between nt −1699 and −2684 (−1699/−2684).

(2) We have mapped two negative regulatory regions in the human *MIP* 5′-flanking sequence at nt −1564/−1696 and −948/−1000. Although the *Alu* elements present at nt −1699/−2684 may contribute to the negative regulation of *MIP*, other sequences immediately downstream may also be essential for this effect.

(3) We demonstrated that the human *MIP* 5′-flanking sequence −253/+42 contains a functional promoter in lens cells but is inactive in cultured non-lens cells, suggesting that this sequence contains regulatory elements responsible for the lens specific expression of *MIP*.

ACKNOWLEDGEMENTS

We would like to thank Dr. Abdul Ally (Biotechnica International, Inc.) and Gabriela Tobal for assistance in dideoxynucleotide sequencing, Marvin Shapiro of the NIH Division of Computer Research and Technology and Dr. David Landsman of the National Library of Medicine's Center for Biotechnology Information for assistance with computer analyses, as well as the National Cancer Institute for allocation of computer time and staff support, in particular Mark Gunnell, at the Advanced Scientific Computing Laboratory of the Frederick Cancer Research Facility, Martin Breitman for the pSVOATCAT plasmid and Graeme Wistow for the critical reading of the manuscript.

REFERENCES

Baeuerle, P.A.: The inducible transcription activator NF-κB: regulation by distinct protein subunits. Biochim. Biophys. Acta 1072 (1991) 63–80.

Borras, T., Peterson, C.A. and Piatigorsky, J.: Evidence for positive and negative regulation in the promoter of the chicken δ1-crystallin gene. Dev. Biol. 127 (1988) 209–219.

Chepelinsky, A.B.: The MIP transmembrane channel gene family. In: Peracchia, C. (Ed.), Handbook of Membrane Channels: Molecular and Cellular Physiology. Academic Press, New York, NY, 1994, pp. 413–432.

Duncan, G. and Jacob, T.J.: The lens as a physiochemical system. In: Davson, H. (Ed.), The Eye, 3rd edition, Vol. 1b. Academic Press, New York, NY, 1984, pp. 159–206.

Ehring, G.R., Zampighi, G.A. and Hall, J.E.: Does MIP play a role in cell-cell communication? Progr. Cell Res. 3 (1993) 143–151.

Englander, E.W., Wolffe, A.P. and Howard, B.H.: Nucleosome interactions with a human *Alu* element. Transcriptional repression and effects of template methylation. J. Biol. Chem. 268 (1993) 19565–19573.

Gorman, C.M., Moffat, L.F. and Howard, B.H.: Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. 2 (1982) 1044–1051.

Harding, J.J. and Crabbe, M.J.C.: The lens: development, proteins, metabolism and cataract. In: Davson, H. (Ed.), The Eye, 3rd ed., Vol. 1. Academic Press, New York, NY, 1984, pp. 207–492.

Howard, B.H. and Sakamoto, K.: *Alu* interspersed repeats: selfish DNA or a functional gene family? New Biol. 2 (1990) 759–770.

Hsieh, C.-L.: Dependence of transcriptional repression on CpG methylation density. Mol. Cell. Biol. 14 (1994) 5487–5494.

Jurka, J. and Milosavljevic, A.: Reconstruction and analysis of human *Alu* genes. J. Mol. Evol. 32 (1991) 105–121.

Kamachi, Y. and Kondoh, H.: Overlapping positive and negative regulatory elements determine lens-specific activity of the δ1-crystallin enhancer. Mol. Cell. Biol. 13 (1993) 5206–5215.

Liu, Q., Tini, M., Tsui, L.-C. and Breitman, M.L.: Interaction of a lens cell transcription factor with the proximal domain of the mouse γF-crystallin promoter. Mol. Cell. Biol. 11 (1991) 1531–1537.

Lok, S., Stevens, W., Breitman, M.L. and Tsui, L.-C.: Multiple regulatory elements of the murine γ2-crystallin promoter. Nucleic Acids Res. 17 (1989) 3563–3582.

Meakin, S.O., Reddan, J.R., Tsui, L.-C. and Breitman, M.L.: A rabbit lens epithelial cell line supports expression of an exogenous crystallin gene characteristic of lens fiber cell differentiation. Exp. Eye Res. 48 (1989) 131–137.

Meehan, R., Lewis, J., Cross, S., Nan, X. and Jeppesen, P.: Transcriptional repression by methylation of CpG. J. Cell Sci. 16 (1992) 9–14 (Suppl.).

Nakamura, T., Donovan, D.M., Hamada, K., Sax, C.M., Norman, B., Flanagan, J.R., Ozato, K., Westphal, H. and Piatigorsky, J.: Regulation of the mouse αA-crystallin gene: isolation of a cDNA encoding a protein that binds to a cis sequence motif shared with the major histocompatibility complex class I gene and other genes. Mol. Cell. Biol. 10 (1990) 3700–3708.

Neumann, J.R., Morency, C.A. and Russian, K.O.: A novel rapid assay for chloramphenicol acetyltransferase gene expression. BioTechniques 5 (1987) 444–447.

Nielsen, D.A., Chou, J., MacKrell, A.J., Casadaban, M.J. and Steiner, D.F.: Expression of a preproinsulin-β-galactosidase gene fusion in mammalian cells. Proc. Natl. Acad. Sci. USA 80 (1983) 5198–5202.

Peek, R., van der Logt, P., Lubsen, N.H. and Schoenmakers, J.G.G.: Tissue-and species-specific promoter elements of rat γ-crystallin genes. Nucleic Acids Res. 18 (1990) 1189–1197.

Peek, R., Niessen, R.W.L.M., Schoenmakers, J.G.G. and Lubsen, N.H.: DNA methylation as a regulatory mechanism in rat γ-crystallin expression. Nucleic Acids Res. 19 (1991) 77–83.

Peek, R., Kraft, H.J., Klok, E.J., Lubsen, N.H. and Schoenmakers, J.G.G.: Activation and repression sequences determine the lens-specific expression of the rat γD-crystallin gene. Nucleic Acids Res. 20 (1992) 4865–4871.

Piatigorsky, J. and Zelenka, P.S. Transcriptional regulation of crystallin genes: cis elements, trans-factors, and signal transduction systems in the lens. Adv. Dev. Biochem. 1 (1992) 211–256.

Pisano, M.M. and Chepelinsky, A.B.: Genomic cloning, complete nucleotide sequence, and structure of the human gene encoding the major intrinsic protein (MIP) of the lens. Genomics 11 (1991) 981–990.

Rogers, L.H.: The origin and evolution of retrotransposons. Int. Rev. Cytol. 93 (1985) 187–279.

Sax, C.M., Cvekl, A., Kantorow, M., Sommer, B., Chepelinsky, A.B. and Piatigorsky, J.: Identification of negative-acting and protein-binding elements in the mouse αA-crystallin −1556/−1165 region. Gene 144 (1994) 163–169.

Yancey, S.B., Koh, K., Chung, J. and Revel, J.P.: Expression of the gene for main intrinsic polypeptide (MIP): Separate spatial distributions of MIP and β-crystallin gene transcripts in rat lens development. J. Cell Biol. 106 (1988) 705–714.