Volume 238, number 2, 253–256 FEB 06370 October 1988

Isolation and analysis of the mouse opsin gene

W. Baehr°, J.D. Falk*°, K. Bugra+, J.T. Triantafyllos+ and J.F. McGinnis+

°*Department of Biological Sciences, Purdue University, West Lafayette, IN 47907* and +*Department of Anatomy and The Mental Retardation Research Center, UCLA School of Medicine, Los Angeles, CA 90024, USA*

Received 5 July 1988; revised version received 15 August 1988

We have identified three overlapping 5'-truncated mouse opsin cDNA clones by immunologically screening a λgt11 retina expression library. Using one of the cDNA clones as a probe, we isolated a 5 kb genomic fragment that encompassed the complete coding sequence for mouse opsin. The coding region for opsin was interrupted by four introns positioned precisely as those previously described for other mammalian opsins. In contrast to the single major opsin mRNA in the bovine and human retina, Northern analysis of mouse retina RNA demonstrated the presence of at least five distinct species of polyadenylated opsin mRNAs. Their sizes ranged from 1.7 kb to 5.1 kb.

Opsin; Visual pigment; cDNA cloning; Gene cloning; Peptide sequence; (Mouse retina)

1. INTRODUCTION

The visual pigment rhodopsin is the primary integral membrane protein found in vertebrate retinal rod outer segments. Rhodopsin is comprised of an apoprotein, opsin, and a 11-*cis* retinal chromophore. Light activation of rhodopsin triggers a cascade of events that result in cell signalling through the hyperpolarization of the rod cell [1,2]. Opsin genes and cDNAs have been isolated and characterized in a number of vertebrate and invertebrate species including bovine, human and *Drosophila* [3]. Regulation of opsin gene expression has been studied extensively in *Drosophila* [4], but not in mammalian systems. The mouse opsin gene is particularly well suited for regulatory studies concerning gene expression. Promoter regions that regulate opsin expression can be examined in tissue culture systems as well as in transgenic mice constructs. Furthermore, mouse strains that carry inherited visual defects may allow us to examine mutations that may affect gene expression. As a first step towards genetic analysis of the mouse phototransduction systems, we have isolated and sequenced a wild-type mouse opsin gene and cDNA, and initiated analysis of transcription of this gene.

Correspondence (present) address: W. Baehr, Cullen Eye Institute, Baylor College of Medicine, Houston, TX 77030, USA

* *Present address:* Eye Research Laboratories, University of Chicago Medical Center, 937 E. 57th Street, Chicago, IL 60637, USA

2. MATERIALS AND METHODS

2.1. *Isolation of RNA*

Total RNA was extracted from eyes (lens removed) or from dissected retinas of adult C57BL/6J mice by homogenization in guanidinium thiocyanate and centrifugation through cesium chloride [5]. Bovine RNA was prepared in a similar manner from freshly dissected retinas. Poly(A) RNA was isolated by two sequential passages through oligo(dT) cellulose columns [6].

2.2. *Isolation of cDNA clones*

We constructed a cDNA expression library in λgt11 [7] according to the method of Young and Davis [8]. Approx. 400000 plaques were screened with antisera generated against total normal adult mouse retinal proteins [9]. Prior to screening, the antisera were absorbed with total retinal proteins from adult photoreceptor-less *rdle* mice to remove antibodies which reacted with antigens common to the retinas of normal and mutant mice. The nitrocellulose filters were probed as described earlier for Western transfers [9]. Four clones (M1–M4) were identified as opsin clones on the basis of their cross-hybridization to a bovine opsin cDNA probe (SP1116, [10]) and

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

their reactivity with a monoclonal antibody (anti-rhodopsin ID4, gift from R. Molday). Restriction of purified λ DNA with *Eco*RI and partial sequencing of cDNA inserts revealed that M1 and M3 were identical.

2.3. *Isolation of genomic fragments*

A wild-type mouse EMBL3 genomic library, prepared from the strain B10 (derived from a C57BL/10 inbred background), was a gift from A.-M. Frischauf and H. Lehrach, EMBL, Heidelberg, FRG. The M2 mouse opsin cDNA insert was labeled by nick-translation and used to screen 160000 plaques under high stringency. Three strongly hybridizing clones were obtained and purified to homogeneity. Isolated λDNA was restricted with *Eco*RI and *Bam*HI, and analyzed by Southern blotting [11]. In all three clones *Eco*RI released a 5 kb fragment that hybridized strongly with the nick-translated M2 insert. *Bam*HI released two fragments, 11 kb and 3 kb. Only the 11 kb fragment hybridized with M2. The positive *Eco*RI fragment was isolated by gel elution and subcloned into pUC13 to yield MOPS1 (fig.1).

2.4. *Deletion subcloning and DNA sequencing*

Inserts of bacteriophage clones were subcloned into M13mp10 and pUC13 vectors, and sequenced with the dideoxynucleotide chain termination method [12]. Random deletion clones generated from M2 by exonuclease III digestion were sequenced directly with the double-stranded plasmid sequencing technique [13]. The genomic fragment MOPS1 was cleaved with *Nco*1, *Dde*1, and *Taq*1, and endlabeled fragments containing exon/intron boundaries were sequenced with the chemical degradation method [14].

2.5. *Northern blotting*

RNA was electrophoretically fractionated on formaldehyde/agarose gels [15] and transferred by capillary action to nylon membranes (Nytran, 0.45 μm, Schleicher and Schuell). The denatured M2 insert was labeled by the random primer method [16] with a kit from Amersham to a specific activity of 2×10^9 cpm/μg. Hybridizations were carried out in 50% formamide/6 × SSC at 42°C. The nylon membranes were washed three times in 0.1 × SSC, 1% SDS, at 50°C for 30 min and exposed to Kodak XAR5 film using an intensifier screen at −70°C.

3. RESULTS AND DISCUSSION

3.1. *Mouse opsin clones*

Screening of the mouse retina cDNA library with polyclonal antisera yielded three truncated mouse opsin clones, M2, M3, and M4 (fig.1). The clone with the longest insert, M2, lacked 5′-untranslated sequences and the first 21 codons of opsin. Southern blotting studies of mouse genomic DNA had indicated that the complete mouse opsin gene was contained as a single copy in a 11 kb *Bam*HI or a 5 kb *Eco*RI fragment [10]. The 5 kb *Eco*RI fragment, MOPS1, was isolated from a genomic EMBL3 clone and shown to contain the complete opsin coding sequence interrupted by four introns, a putative transcriptional starting point, and a polyadenylation signal (fig.1).

3.2. *Mouse opsin gene structure*

The transcription start point was tentatively assigned (position 1, fig.2) based on a TATAA box at position −26 to −30 [17], and nucleotide sequence homology to the bovine and human genes [18,19]. Comparison of the region 5′ to the transcription start point showed that the −8 to −80 regions are highly conserved in the three known mammalian opsin genes. However, unlike the bovine and human opsin genes, the mouse opsin gene had no CAAT box 70–80 bp upstream of the transcription start point. Such a sequence was identified 30 bases further upstream at position −123 to −119 (fig.2). The length of the exons (hatched boxes in fig.1) and the positions of the introns were identical to those described for bovine and human opsins [18,19]. The lengths of the introns varied considerably. The human and mouse introns are, 1.78 kb and 1.38 kb for intron 1; 1.2 kb and 1.08 kb for intron 2; 116 bp and 117 bp for intron 3; 833 bp and 920 bp for intron 4, respectively.

3.3. *Gene transcripts*

High stringency Northern blot analysis showed that the mouse eye contained five major species of opsin mRNA with sizes of 5.1 kb (5%), 3.9 kb (10%), 3.1 kb (20%), 2.2 kb (35%), and 1.7 kb (30%) (fig.3). None of these bands were detected in RNA isolated from mouse brain (fig.3), mouse liver, or the photoreceptorless retinas of adult mutant (*rd*) mice (not shown). The mRNA in each of the bands appeared to be polyadenylated because each bound to and could be recovered from an oligo(dT) cellulose column. The identification of 5 transcripts is in agreement with data previously presented [20]. Bovine retina RNA contained only one major species (95–98%) of opsin mRNA (fig.3) with a size comparable to the 3.1 kb mouse transcript.

The distance between the putative transcription start and a polyadenylation signal (excluding introns) on the mouse opsin gene (fig.1) is 1487 bases. The addition of a poly(A) tail would result in an approx. 1.7 kb processed transcript which may correspond to the smallest transcript in fig.3.

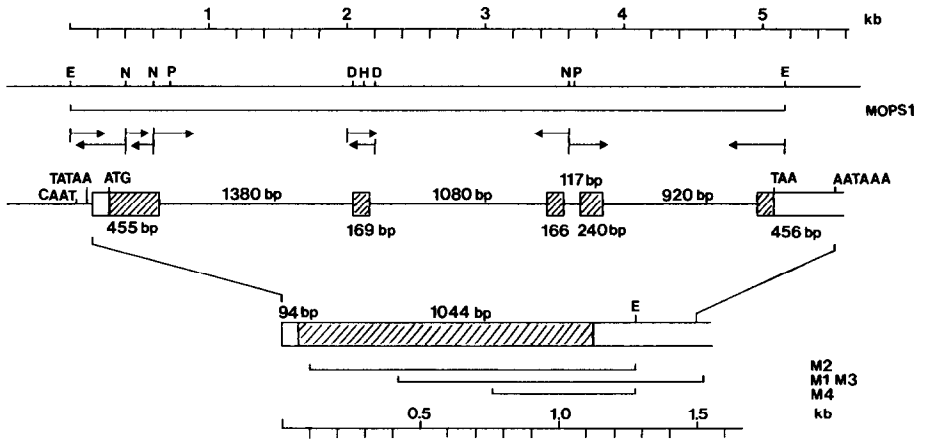

Fig.1. Restriction map of mouse opsin clones. The top line depicts a restriction map of the 5 kb *Eco*RI fragment that contains the mouse opsin gene. Single letters symbolize restriction sites that were used for sequencing of endlabeled fragments [14] (half arrows): E, *Eco*RI; N, *Nco*I; T, *Taq*I; P, *Pst*I; D, *Dde*I; H, *Hae*II. Hatched boxes underneath the restriction map indicate the extent of exons, and lines connecting the exon boxes denote the length of introns. The exons are graphically spliced to a model symbolizing mouse opsin mRNA. M3, M3, and M4 are inserts of cDNA λgt11 clones. Arrows give the direction of DNA sequencing.

```
                                                                    GAATTCCCAAGAGGACTCTGGGGCAGACAAGATGAGACACCCTTTCCTTTCTTTACCTAAGGGCCTCCACCCGAT -147
GTCACCTTGGCCCCTCTGCAAGCCAATTAGGCCCCGGTGGCAGCAGTGGGATTAGCGTTAGTATGATATCTCGCGGATGCTGAATCAGCCTCTGGCTTAGGGAGAGAAGGTCACTTTATA -27

AGGGTCTGGGGGGGTCAGTTACCTGGAGTTGCGCTGTGGGAGCCGTCAGTGGCTGAGCTCGCCAAGCAGCCTTGGTCTCTGTCTACGAAGATCCGTGGGGCAGCCTCGAGAGCCGCAGCC  94

ATGAACGGCACAGAGGGCCCCAATTTTTATGTGCCCTTCTCCAACGTCACAGGCGTGGGTCGGAGCCCCTTCGAGCAGCCGCAGTACTACCTGGCGGAACCATGGCAGTTCTCCATGCTG 214
 M  N  G  T  E  G  P  N  F  Y  V  P  F  S  N  V  T  G  V  G  R  S  P  F  E  Q  P  Q  Y  Y  L  A  E  P  W  Q  F  S |M  L  40
                                                  A        V              Y

GCAGCGTACATGTTCCTGCTCATCGTGCTGGGCTTCCCCATCAACTTCCTCACGCTCTACGTCACCGTACAGCACAAGAAGCTGCGCACACCCCTCAACTACATCCTGCTCAACTTGGCC 334
 A  A  Y  M  F  L  L  I  V  L  G  F  P  I  N  F  L  T  L  Y  V| T  V  Q  H  K  K  L  R  T  P  L |N  Y  I  L  L  N  L  A  80
 ——————————— helix 1 ———————————————————————————————————————————————————————————————————————— helix 2 ————————

GTGGCTGACCTCTTCATGGTCTTCGGAGGATTCACCACCACCCTCTACACATCACTCCATGGCTACTTCGTCTTTGGGCCCACAGGCTGTAATCTCGAGGGCTTCTTTGCCACACTTGGA 454
 V  A  D  L  F  M  V  F  G  G  F  T  T  T  L  Y  T| S  L  H  G  Y  F  V  F  G  P  T  G  C  N  L  E  G |F  F  A  T  L  G  120
 ——————————————————L———————————S———————————————————————————————————————————————————————————————————————————————

          intron 1
Ggtatgagcaga.....gaactgccttacagGTGAAATCGCCCTGTGGTCCCTGGTGGTCCTGGCCATTGAGCGCTACGTGGTGGTCTGCAAGCCGATGAGCAACTTCCGCTTCGGG 541
                                G  E  I  A  L  W  S  L  V  V  L  A  I  E  R  Y  V  V  V  C| K  P  M  S  N  F  R  F  G  149
                                ——————————————————————————————————— helix 3 ——————————————————————————————————————

                                                                                                    intron 2
GAGAATCACGCTATCATGGGTGTGGTCTTCACCTGGATCATGGCGTTGGCCTGTGCTGCTCCCCCACTCGTTGGCTGGTCCAGgtaaggctcag.....gggcagccacagGTACATC 631
 E  N  H |A  I  M  G  V  V  F  T  W  I  M  A  L  A  C  A  A  P  P  L  V  G  W  S| R                                  Y  I  179
         ————————A————————V——————— helix 4 ———————————A————————

CCTGAGGGCATGCAATGTTCATGCGGGATTGACTACTACACACTCAAGCCTGAGGTCAACAACGAATCCTTTGTCATCTACATGTTCGTGGTCCACTTCACCATTCCTATGATCGTCATC 751
 P  E  G  M  Q  C  S  C  G  I  D  Y  Y  T  L  K  P  E  V  N  N  E |S  F  V  I  Y  M  F  V  V  H  F  T  I  P  M  I  V  I  219
            L                                                      ——————————————————————— helix 5 ———————————————I————

                                                intron 3
TTCTTCTGCTATGGGCAGCTGGTCTTCACAGTCAAGGAGgtatgagcagg.....ctgtgtccctgcagGCGGCTGCCCAGCAGCAGGAGTCAGCCACCACTCAGAAGGCAGAGAAGGAA 841
 F  F  C  Y  G  Q  L  V  F  T  V| K  E                                  A  A  A  Q  Q  Q  E  S  A  T  T  Q  K  A  E  K  E  249
 ——————————————————————————————|

GTCACCCGCATGGTTATCATCATGGTCATCTTCTTCCTGATCTGCTGGCTTCCCTACGCCAGTGTGGCCTTCTACATCTTCACCCACCAGGGCTCCAACTTCGGCCCCATCTTCATGACT 961
 V  T  R |M  V  I  I  M  V  I  F  F  L  I  C  W  L  P  Y  A  S  V  A  F  Y  I  F  T| H  Q  G  S  N |F  G  P  I  F  M  T  289
         ————————A————————V——————— helix 6 ———————————| ——————————————————————————————————————————————————————
                                                                                      intron 4
CTGCCAGCTTTCTTTGCTAAGAGCTCTTCCATCTATAACCCGGTCATCTACATCATGTTGAACAAGCAGgtgttcaagcagagc.....tgccttccagTTCCGGAACTGTATGCTC 1048
 L  P  A  F  F  A  K  S  S  S  I  Y  N  P  V  I  Y  I  M  L  N| K  Q                                  F  R  N  C  M  L  318
-I————————————————A—A———— helix 7 ———————————M————|

ACCACGCTGTGCTGCGGCAAGAATCCACTGGGAGATGACGACGCCTCTGCCACCGCTTCCAAGACGGAGACCAGCCAGGTGGCTCCAGCCTAAGCCTGGCCAGAGACTGTGGCTTGAAAG 1168
 T  T  L  C  C  G  K  N  P  L  G  D  D  D  A  S  A  T  A  S  K  T  E  T  S  Q  V  A  P  A  *                              348
    I                                  E           V

TAGGAGTCTCCTGTCCCCACACACCCCAGCCACAGCCCCCACCAGGAGCAGCACCGTTGGGATGAGGTCATGCAGGCTCCCTCAGTGTTCTTTTCTTTGTTTTTAATGAATTCATGAAAG 1288
CAAAATGAGGCTCCCCACTCAATGGGGACAGCTTGACAAAGGGCATCCATCCACCAAGACCATCCTCAACCTGGAGTCCCCAATTCCCGGGGGGCCAGCGGGATCTGTACCCCTCCCTCA 1408
GCTTGTCTATCAGGAACATGACAAGTGTCCCGGCTTAGGGCTAAATGTCTAGGACAGAATGGAACACATAGTAGCTGAATTAATAAATGCTAGCTTGGATGAAGGGAGGA          1518
```

Fig.2. Sequence of the mouse opsin gene. The nucleotide sequence is a composite of genomic and cDNA sequencing. Position 1 is the presumptive transcriptional start point. The underlined sequences upstream of this point are the TATAA and CAAT boxes. Intervening sequences within the 1044 coding region are depicted in lower case letters (only the first 9–14 bases after or before each exon/intron junction are shown). In the 3′ flanking sequence, the AATAAA polyadenylation signal is underlined.

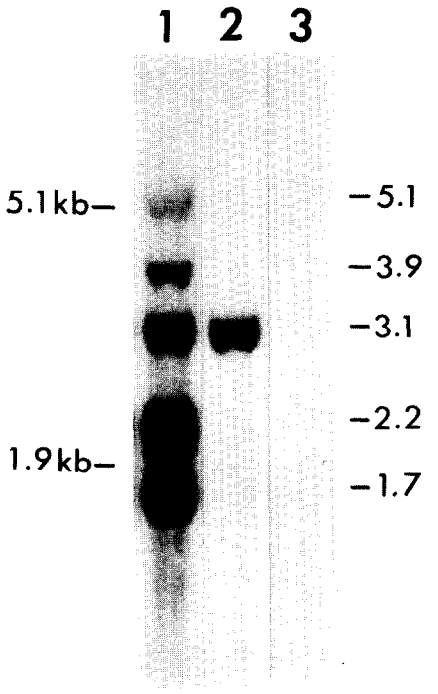

Fig.3. Northern analysis of opsin mRNA. 5 μg of total RNA was loaded per lane (lanes: 1, mouse retina; 2, bovine retina; 3, mouse brain). All samples were treated identically. However, in the composite print shown, lane 1 is the result of an hour exposure; lane 2 a 12-h exposure; and lane 3 a 24-h exposure. The numbers on the left are the sizes (kb) of mouse ribosomal RNA, and the numbers on the right are the sizes of the mouse opsin mRNAs determined by comparison with an RNA ladder.

A completely unprocessed transcript should result in a 5.2 kb species which correlates with the largest observed band in fig.3. Two polyadenylation signals in the 3′-flanking region have been reported for the human opsin gene, and the second site is inferred for the bovine gene resulting in the major 3.1 kb mRNA. The mouse opsin gene may have a similar second site. We are currently investigating the precise nature of the different RNA species with intron and exon specific probes.

3.4. *Mouse opsin amino acid sequence*

The length of 348 amino acid residues appears to be conserved for all mammalian rod opsins. There are 24 substitutions in mouse opsin, as compared to bovine opsin (fig.2), and 19 substitutions as compared with human. Nine of the residues are variant in all three opsin sequences. Half of the mouse substitutions occur in the transmembrane hydrophobic domains (fig.2). Most of these substitutions are conservative and none of them affect the charge distribution near the retinal chromophore. No changes were observed in the sequence of the cytoplasmic helix-connecting loops in these mammalian rod opsins, emphasizing the importance of these loops as points of contact with the photoreceptor cell G protein.

Acknowledgements: We extend our appreciation to M.L. Applebury, in whose laboratory this work was carried out. We thank Kristi Volpp for excellent technical assistance during this work. This research was funded by grants from the National Eye Institute, EY04801 (M.L. Applebury). Core Grant EY02723 (Purdue University) and Training Grant support from EY07008 to J.D.F. The work at UCLA was funded by grants EY6085 and EY6639 to J.F.M. and fellowships from the National Retinitis Pigmentosa Foundation and Fight for Sight PD87-054 (K.B.). J.T.T. was a student fellow of the Fight for Sight Foundation.

REFERENCES

[1] Hargrave, P.A. (1982) Progr. Ret. Res. 1, 1–51.
[2] Applebury, M.L. and Hargrave, P.A. (1986) Vision Res. 26, 1881–1895.
[3] Falk, J. and Applebury, M.L. (1987) Progr. Ret. Res. 7, 89–109.
[4] Zuker, C.S., Mismer, D., Hardy, R. and Rubin, G.M. (1988) Cell 55, 475–482.
[5] Chirgwin, J.M., Przybala, A.E., McDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[6] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1409–1412.
[7] McGinnis, J.F. and Leveille, P.J. (1986) J. Neurosci. Res. 16, 157–165.
[8] Huynh, T., Young, R.A. and Davis, R.W. (1985) in: DNA Cloning, A Practical Approach (Glover, D.M. ed.) vol.1, pp.49–78, IRL, Oxford.
[9] McGinnis, J.F. and Leveille, P.J. (1985) Curr. Eye Res. 4, 1127–1135.
[10] Martin, R.L., Wood, C., Baehr, W. and Applebury, M.L. (1986) Science 232, 1266–1269.
[11] Southern, E. (1975) J. Mol. Biol. 98, 503–517.
[12] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Chen, E.Y. and Seeburg, P.H. (1985) DNA 4, 165–170.
[14] Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[15] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[16] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[17] Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383.
[18] Nathans, J. and Hogness, D.S. (1983) Cell 34, 807–814.
[19] Nathans, J. and Hogness, D.S. (1984) Proc. Natl. Acad. Sci. USA 81, 4851–4855.
[20] Bowes, C. and Farber, D. (1987) Exp. Eye Res. 45, 467–480.