Molecular Vision **1998; 4**:14 <http://www.molvis.org/molvis/v4/p14>
Received 7 August 1998 | Accepted 24 August 1998 | Published 4 September 1998

Molecular Characterization of the Mouse Gene Encoding Cellular Retinaldehyde-binding Protein

Breandán N. Kennedy,[1] Jing Huang,[2] John C. Saari,[2] John W. Crabb[1]

[1]Adirondack Biomedical Research Institute, Lake Placid, NY, 12946; [2]Departments of Ophthalmology and Biochemistry, University of Washington School of Medicine, Seattle, WA, 98195

Purpose: To clone and characterize the mouse gene encoding cellular retinaldehyde-binding protein (CRALBP). CRALBP appears to modulate enzymatic generation and processing of 11-*cis*-retinol and regeneration of visual pigment in the vertebrate visual cycle. Mutations in human CRALBP segregate with autosomal recessive retinitis pigmentosa.

Methods: A genomic clone encompassing the 5' end of the CRALBP gene through exon 6 was isolated from a mouse 129/Sv genomic DNA library. Exons 7 and 8 were PCR amplified from mouse eye cDNA and 129/SvJ genomic DNA. The gene structure was determined by automated DNA sequence analysis.

Results: The sequence of 6855 nucleotides was determined, including all 8 exons, 3 introns plus 3932 and 629 bases from the 5'- and 3'-flanking regions, respectively. The lengths of introns 3-6 were determined by PCR amplification. Northern analysis identifies a ~2.1 kb transcript in mouse eye; Southern analysis supports a single copy gene.

Conclusions: The mouse CRALBP gene is similar to the human gene; the coding sequence is ~87% identical, the non-coding sequence ~65% identical. In contrast to the human gene, the mouse gene contains a consensus TATA box. One of two photoreceptor consensus elements important for CRALBP expression in human retinal pigment epithelium is also present in the mouse gene. Additional conserved and species-specific consensus sequences are identified. The mouse CRALBP genomic clones and structure provide valuable tools for developing an in vivo model to study protein function and gene regulation.

Cellular retinaldehyde-binding protein (CRALBP) serves as a substrate carrier protein for enzymes of the mammalian visual cycle in vitro, modulating whether 11-*cis*-retinol is stored as an ester in the retinal pigment epithelium (RPE) or oxidized by 11-*cis*-retinol dehydrogenase to 11-*cis*-retinal for visual pigment regeneration [1]. Recent evidence suggests that CRALBP also stimulates the isomerohydrolase responsible for generation of 11-*cis*-retinol [2]. CRALBP is strongly expressed in RPE and Müller cells of the retina, where the binding protein carries endogenous 11-*cis*-retinol and/or 11-*cis*-retinal [3]. In addition, the protein is expressed in ciliary body, cornea, pineal gland, optic nerve, brain, and transiently in iris. In at least brain and optic nerve, CRALBP lacks 11-*cis*-retinoid ligands and apparently serves functions unrelated to visual pigment regeneration [4]. Notably, a missense mutation in human CRALBP that destroys retinoid-binding capability has been genetically linked with autosomal recessive retinitis pigmentosa [5].

RLBP1, the gene encoding human CRALBP, has been localized to human chromosome 15q26 [6] and its structure reported [7]. Previously, we demonstrated that *RLBP1* domains -2089 to -1539 bp, -243 to +80 bp, and two photoreceptor consensus elements (PCE1) between *RLBP1* positions -165 to -140 direct CRALBP expression in human RPE cultures but not in non-ocular cell cultures [8]. As part of ongoing efforts to develop an in vivo model for studying CRALBP function and gene regulation in transgenic mice, we have cloned and characterized the mouse CRALBP gene (mouse *Rlbp1*). Here we report over 6 kb of *Rlbp1* sequence, including 5' and 3' flanking regions and the entire coding sequence. Conserved and species-specific consensus sites of potential regulatory elements are identified and the deduced mouse CRALBP protein structure discussed.

*To whom correspondence should be addressed: Dr. John W. Crabb, Protein Chemistry, Eye Institute, Cleveland Clinic Foundation, 9600 Euclid Avenue, Cleveland, OH 44195; Phone (216) 444-5832; email: crabbj@cesmtp.ccf.org

Dr. Kennedy's present address is: Center for Transgene Research, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, 46556.

METHODS

Materials and Reagents—A female mouse 129/Sv genomic library in lambda FIX II (Stratagene, Inc., La Jolla, CA) was used for cloning part of the mouse CRALBP gene. C57BL6 mice were sacrificed by cervical dislocation and eyes dissected and frozen in dry ice/ethanol within 60 s of death [9]. 129/SvJ genomic DNA was purchased from Jackson Labs, Bar Harbor, ME. RNA and DNA molecular weight markers were from Novagen, Inc. (Madison, WI) and Life Technologies, Inc. (Gaithersburg, MD).

Mouse Genomic Library Screening and Characterization of the Mouse CRALBP Gene— Approximately 1 X 10^7 recombinant phage from the 129/Sv mouse genomic DNA library (1.0 X 10^{10} pfu/ml) were plated on *Escherichia coli* strain BB4(LE392). Plaques were screened with human (1317 bp) and bovine (1173 bp) CRALBP cDNA probes [10]. Several positive clones were plaque-purified and one, designated λA2, was selected for further analysis. Plasmid subclones mp3.2SK and mp2.8SK were created from λA2 by *EcoR1* restriction digestion and subclone mp3.6CRII by the polymerase chain reaction (PCR). 129/SvJ mouse genomic DNA was PCR

amplified using primers designed to mouse and human CRALBP and 10-100 ng of genomic DNA or plasmid DNA as template. Standard PCR cycle conditions were 94 °C/30-45 s, 55-63 °C/30-45 s and 72 °C/min per kb of expected product for 35 cycles. The amplification systems used were *Taq* polymerase (Fisher, Pittsburgh, PA) for plasmid DNA and the Expand Long Template PCR system (Boehringer Mannheim, Indianapolis, IN) for genomic DNA templates.

Mouse Eye cDNA Preparation and Analysis of CRALBP cDNA— RNA was isolated from C57BL6 mouse eyes with the RNA Ultraspec II Isolation system (Cinna-Biotecx, Houston, TX). First strand cDNA synthesis was performed according to the Superscript preamplification System (Life Technologies) using primer NNT_{20}(N=G,A,T or C), a modified oligo-dT primer with elevated melting temperature. Mouse CRALBP cDNA was amplified with primers specific to mouse or human CRALBP using the standard PCR conditions. To completely sequence mouse CRALBP exons 7-8, nested 5' primers were successively used in combination with primer NNT_{20} to specifically amplify mouse C57BL6 eye cDNA. Aliquots of synthesis products were serially diluted 1:20 in consecutive Touchdown-PCR reactions [11]. Touchdown-PCR cycle conditions were 94 °C/30 s, 55 °C/30 s and 72 °C/2 min for 10 cycles followed by 94 °C/30 s, 48 °C/30 s and 72 °C/2 min for 20 cycles.

Northern Blot Analysis— Total RNA was isolated from mouse eyes, liver and kidney, electrophoresed on 1% agarose gels containing 6.6% formaldehyde, blotted to Zeta-Probe GT membranes (Bio-Rad, Hercules, CA) in 20 x SSC, crosslinked to the membrane by UV irradiation (Stratalinker, Stratagene, Inc.) and hybridized with mouse CRALBP cDNA probes. Blocking and hybridization were carried out in the presence of Hybrisol II (Oncor, Gaithersburg, MD) and 20% formamide at 52 °C; final high stringency washing was at 63 °C with 0.1X SSC/0.1% SDS. Two mouse CRALBP cDNA probes (240 bp and 290 bp) were used in Northern analyses. The 240 bp probe spanned sequence in exons 5 and 6 and was generated by PCR with forward primer (5'-AGGCTCTCCGCT-GCACTATC-3') and reverse primer (5'-GCATGTCC-ACCATCTTCTTGAG-3'). The 290 bp probe spanned sequence in exons 6 and 7 and was generated by PCR using forward primer (5'-GGAGAAACTGCTGGAAACTGAGG-3') and reverse primer (5'-CAAGAAGGGCTTGA-CCACATTGTAGG-3').

Southern Blot Analysis— Mouse genomic DNA samples were subjected to agarose electrophoresis using Tris acetate buffers and transferred to Zeta-Probe GT positively charged nylon membranes (Bio-Rad) by downward alkaline capillary transfer [12]. A 230 bp probe spanning the exon/intron 5 boundary was generated by PCR with forward primer 5'-AGGCTCTCCGCTGCACTATC-3', reverse primer 5'-TGCTGGAAAGATGCTGACTACC-3'. In a modification of the random-primer labeling protocol [13] primers specific to the probe were substituted for random primers (Specific-primer labeling). The labeled probe was purified from unincorporated nucleotides using Sephadex G-25 chromatography [13]. The denatured probe was incubated with the membrane in hybridization solution (6X SSC, 5X Denharts buffer, 0.5% SDS and 1% sheared salmon sperm DNA) at 68 °C overnight. Membranes were washed under moderate stringency conditions (0.2X SSC/ 0.1% SDS at 42 °C) prior to autoradiography.

DNA Sequencing and Oligonucleotide Synthesis— Oligonucleotide primers were designed using PRIMER DESIGNER version 2.0 software (Scientific and Educational Software). Plasmid DNA templates for sequencing were

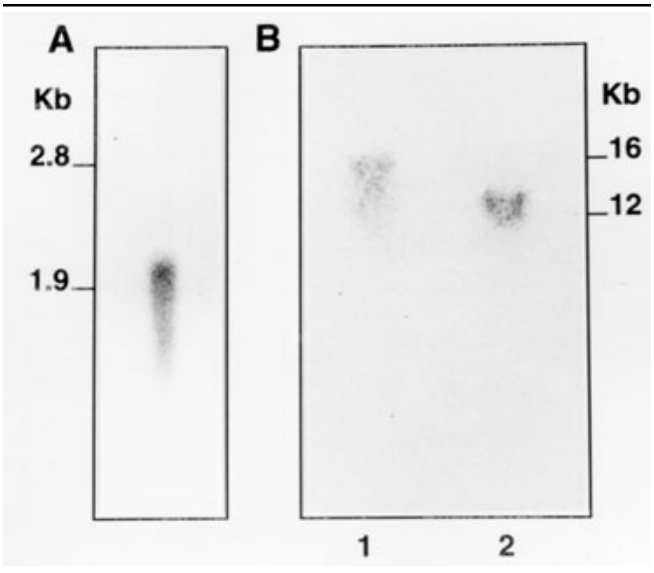

Figure 1. Northern and Southern analyses. (A) Northern blot. 10 µg of total mouse eye RNA was electrophoresed on a 1% agarose-formaldehyde gel, blotted to nylon membrane and hybridized with mouse CRALBP cDNA probes as described in Materials and Methods. The indicated sizes (in kilobases) are estimated using RNA molecular weight markers. (B) Southern blot. 10 µg of mouse genomic DNA was digested with *BamHI* plus *XhoI* (lane 1) or *BamHI* plus *KpnI* (lane 2), electrophoresed on a 1% agarose gel, blotted to nylon membrane and hybridized with a 230 bp mouse CRALBP genomic probe as described in Materials and Methods. The indicated sizes (in kilobases) are estimated using DNA molecular weight markers.

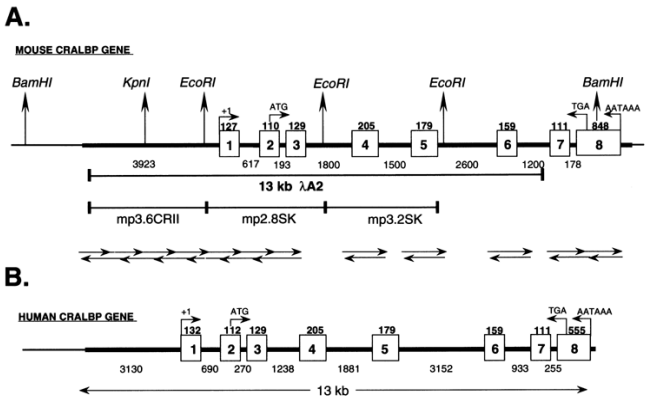

Figure 2. Organization of the Mouse CRALBP Gene. Shown are (A) a restriction map and the clones used in the mouse CRALBP gene sequence determination, and (B) a comparison with the human CRALBP gene. The length of exons (boxed), introns and the flanking regions are indicated (in bases). Horizontal arrows illustrate the regions of the gene that were sequenced. Sequence down stream of exon 6 was obtained by direct analysis of PCR products from mouse cDNA and 129/SvJ genomic DNA as described in the text.

purified using the Wizard Mini-Preps System (Promega, Madison, WI). For direct sequencing, PCR products were isolated from Sea Plaque low melting point agarose (FMC Bioproducts, Rockland, ME) and purified with the Wizard PCR Preps System (Promega). DNA Sequencing was performed using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer, Foster City, CA) with AmpliTaq DNA Polymerase, FS. Synthesis of oligonucleotides and collection of automated DNA sequencing results were performed in the Molecular Biology Core Facility, Adirondack Biomedical Research Institute, using Perkin Elmer, Applied Biosystems Division instrumentation and reagents (model 392 DNA/RNA Synthesizer, model 373, DNA Sequencer). All mouse CRALBP DNA sequences were

```
                  cc ctttgctagc actgagaccc tttaatacgc -3901
ctcctcatgt tgtggggacc ccttaaccat aaaattattt ttgttgctcc -3851
atcataacta taatttttgtc accttttatga atcataatgt aaatatctga -3801
tattcaggat atctgatatg caactcttgt gaaagggtcc atcaaccacc -3751
ccatccccca ccctcagtga tcatgaccca caggttgggt ctaggcagcc -3701
ctctcctgta agactctggc cagccttagc ttacagggga ttccctgagt -3651
gaacagcaca gagccgaaaa tcgatgcaaa aagcaagagg aatttattgc -3601
tccaacatgt tggggtcgcc ctgcacatga agagagagac cgaccacagg -3551
cagcccttcc tttacagctt ctccaggcag cagccattag gcacaatatg -3501
attggcagaa cagtatgact tttaaactga ttggtctata gggaatgagg -3451
taggtggcca acagtatctc tgagcctccc ctacccgccc gcaggtggcc -3401
agtggttggt ccctcctccc ggtctgagga atgtaattag ccccttcctt -3351
ccggagggga tgggtgcgtt ctttaagacc tttctcttcc aggaggggag -3301
gggtctggtg gaattttcta aagtacctga gctgacctct tcactccaac -3251
tagaccatgt catggtagtc ctcttggtac acagggatgc ttggtgatga -3201
atcacagagc ctcttagcag ctgggactct tacctcccat gagcatcaca -3151
gccaggagag ttctttgggc ataacagcat gccacagcct aggtgaatta -3101
accaatagaa aatcattctc tcggaatctc ggtagttgga ggtctaagat -3051
caaggtatta tcctttggtg tcacctcctt atcttataga taaaccctga -3001
actgtggctg gatttcagag aacgacctgt gcacatgagc tcaaccctac -2951
cctgagcaca tggaggagag agaaccagga ccaaaggtgg gctgcttggg -2901
tccagactga agatgaattt aattacagac aactatctgt ggtggaagta -2851
gggatctccc ggacattgga ggccccaca gagatctaaa ggatgtgtat -2801
ccctcatacc ccggtgataa accaaggcag agtctgggtg ctcctttgac -2751
tgttggatgc attcccttag cctgggtgct cactggctct ccattcactg -2701
ggtacaactc agaactcccc agaatacttc ccagggattg ggcagctaag -2651
cccctgaagt gggtactttg catctcagga ctttccccac agcttcagct -2601
ggctctctag tcatcccagt tctcagaggg gactgagtca tgtcgtggtg -2551
acataaccac actgaaaaaa gtgcagtagc aaggacagaa agtcggacaa -2501
gggagatctg ctgaggcccc acacccatcc cctgcttcct tgcctttgcg -2451
aggtagagat cggcagagat cggagggtca gggtacctag caagtcctcc -2401
gctgctcagc ctcgcccgct ccatttccct ctgccacatg ggacagtgga -2351
ccataggcct gggttgcagc aatgatcaca tttataattt ctcttcactt -2301
ctctatatgt cctgcccttt gtgaagaagg agtagtctcc caggcccccc -2251
tgcctcccag tttgagaagc gagatggcag ggaaatgaga cggggctggc -2201
ccagatccag tgaagtggct tccacatggc tgaacagaga gctccaagca -2151
ggagctcaca aagcagggct ttgatggtag ccccgcaggc cgggcctctc -2101
ccctctccct cagggtgctt cacttgcttc attctgggcc tcctgtgagc -2051
ctgactgaat gggaaacggt gggcctgaga agttccttgt gacacactaa -2001
tcccagcttg ctgacctgac cacaccccca aagaaagcag cctcccgagc -1951
tccagaacaa ttcaggcatc cagtagcccc caagaaagag ctgtcaggta -1901
agtcagtaac aaggcctggc tgggccagtg agtataagaa agagtgtgtg -1851
catgggagcg gggtaaatga tcagggatgg gtggcgaaca ccacgtctca -1801
gattgcttgg accggcctgt aaaggcgtac gccatttccg cagggtgaga -1751
cagaccgatc cacgggcaga cgcaacacat acgaatgtgg attacctaac -1701
tcaattagaa tggcaggtga tcaaatactc tttatgccct agtattggta -1651
attaattgta ctttatttgc taattctaaa tgccggagaa catcgatgtg -1601
gagatagggg tagcatgttt ttattacata catgcctaca gctaatgggt -1551
aagaacattt ttaaaataag agataagtgc ttctccgtgc tgacgttaat -1501
cttgcacttc taagatctgg gacttgggct gctgatctgc acatttggct -1451
agggataatg cactaatacc tctcgagtct tgctttgctt tgtttgcttg -1401
aggcagggtt tctctacgta gccctagcta tcctggaact cactatgaag -1351
aacagactgg cttcaaaatc agagattccc ctagacactg ggattaaaga -1301
tgcgtaccac tatacccagc ttttgctttt cgtttttttg agatagagtc -1251
tcatgtagca caggctggtt ttgaattaac tatgtagcca aggataacct -1201
taactaagga ttctcttgcc tctatctttt tttttctttc ctttttaaaa -1151
attactagat attttctttg tttacatttc aaatgttatc gcctttccta -1101
gtttcctctc caaaagtccc ctataccctc acccccgccc tgctccccaa -1051
cccacccatt tctgtttcct ggccctggca ttcccctgta ctggggcata -1001
taatcttcgc aagactaagg gcctctcctc ccattgatgg ccaactaggc  -951
catcttctgc tacataagca gctagagaca caagctctgg ggggtactgg  -901
ttagttcata ttgttgttcc tcctataggg ctgcagaccc cttcagctcc  -851
ttgggtactt tctctagctc cttcgttggg gaccctgtgt tccatcctat  -801
agatgactgt gagcgtccac ttctgtattt gccaggcact gcttttgcct  -751
ctatcttgca agttctggga atacagatgt gtgccacaac agtcaactct  -701
tggtctttac acccacatat ctggatgtct gtctactatc caaagagatg  -651
gcttgtactt gattcttgtg ggtagatccc agcccccaag aattccctgg  -601
tcctttgggg gctgtggtag ggaagggata ggccctcaga tattctctga  -551
gatcctcaag tttcagtgta caaattgggt taatctcagt gctagcaaag  -501
ggagaaaaga ctgttgtcga tcaggagtcc tctctcggtt atgtcacaga  -451
accctgtggt cctcacagcg gtcacttttg gtagatatta ttatccttcg  -401
ttggctctca atctcagaag gattaaatac cttgcctatg gtcacacagg  -351
tggtaaaagt cagattcaaa ttcagatctc tctaacctat aaggtttctg  -301
tcaccagctc ccccagggac tttgcaaaga gatgacttgg cttgaggcaa  -251
agagacttag tgtcaggtgg atacaggttc aagtcctagc ttaaatacca  -201
actatgcgta agtgaccgag ctgaccaatg gaacctagcc acaggcagaa  -151
acagaatgta aagatgttca cagtcactgg ggacatttag gtcttggagg  -101
agaaaatctg gaaatgtggg tcatgatgga tatctgttct ggtggcccag   -51
aaatctatgc tggtgtatat tatctaatta catattaatc cttctatccc    -1
Exon 1
AACAAGGGAC GGTCAAGACT CGGATGAAAT CACACAACTT GCAAGTGTGA   50
GAGACAGCAT TGCAATCCAG GCCCCACACC CTGCTGCCCC AAGCTGCTTC  100
TTGGTGTCTA CAGAAGTCAC AACTTAGgta agtatccttg ggctctgtga  150

ccttggacca gtcacttact ctcctggccc tgttttcccc ttaaataaaa  200
caggagtgga gctgccatgc tcgtgtctga gttaacgaat ttagaagtcc  250
ttagaatgtt gtaaaccctct gtaacaaaga cccaggtctg attatataac  300
tgctcaccat ggctcctaag aatgaatgcc ctgggtatgg agccatccac  350
taggtggaga ggttttctggg gtgcacattt gtactgggga agcaaggcta  400
cagagaagaa gagagattgc cttacactgc agtgagctga ccttgcagct  450
gagaatgggg cccaattctg gtttgttttt ggttagccac agctgaagac  500
tacttagtgt attggtagta tgctccctcc aaatgcagtg ttcccaaagt  550
atagcccagg aaacaagtcc ttcagagcta ggtagggaag gagggctcta  600
attttagtcc tctggttaga ctcacagggg ccaacactcg tctatatttc  650
tgaccaatca gctagctgtt gggagggact tgaatgcttg agatccactg  700
Exon 2
ttctgagaca aacacaaaat agtttctcta atatcactgc agTCTTGACT  750
TCTTTGACTG CTGGACAAAG GCTACCCGGA GAGACCAGGG CAGAGCAAAC  800
AGGTTCTTTC TCTCCCCCCT TCAGTGTCTT CAAGAGGCAG TATGTCAGAC  850
GGGgtgagta tcttgggctc ttcatccctt cttcccttcc tccccacacc  900
tcaacaatct ctctggggcc ccgacctacc aaagggagga gactgagcac  950
agatcttggc caaaggcact gtgcaaaccc atttgccagt tgttctggct 1000
Exon 3
gccaagggct gagctgtccc accttctatg actcctttcc gaccagGTGG 1050
GCACTTTCCG CATGGTTCCT GAAGAGGAGC AGGAGCTCCG AGCACAACTG 1100
GAGCAGCTCA CAACCAAGGA TCATGGTCCT GTCTTTGGCC CATGCAGCCA 1150
GCTGCCCCGC CACACTTTGC AGAAGgtgag gtggc..... .......... 1185
Exon 4
cgtgttctag GCCAAGGATG AGCTGAATGA AAAGGAGGAG ACCCGGGAGG
AAGCGGTGAG GGAGCTACAG GAGCTGGTAC AGGCACAGGC AGCTTCTGGC
GAGGAATTGG CCCTGGCAGT GGCTGAGAGG GTGCAGGCAA GAGACAGCGC
CTTCCTCCTG CGCTTCATCC GTGCCCGCAA GTTCGATGTG GGTCGTGCTT
ATGAGCTGCT CAAAGgtaag gcct..... .......... ..........
.......... .......... .......... .......... ..........
Exon 5
ctccttccag GCTATGTGAA CTTCCGCCTC CAGTACCCTG AACTCTTCGA
TAGTCTCTCC ATGGAGGCTC TCCGCTGCAC TATCGAGGCC GGATACCCTG
GTGTCCTTTC CAGTCGGGAC AAGTATGGTC GAGTGGTTAT GCTCTTCAAC
ATCGAAAACT GGCACTGTGA AGAAGTGACC TTTGATGAGg tcagccgg..
Exon 6
ATCTTACAGG CATATTGTTT CATTTTGGAG AAACTGCTGG AAAATGAGGA
AACCCAAATC AACGGCTTCT GTATTGTTGA GAACTTCAAG GGCTTCACCA
TGCAGCAGGC AGCAGGGCTC CGCCCCTCGG ATCTCAAGAA GATGGTGGAC
ATGCTCCAGg tgaggcctc. .......... .......... ..........
Exon 7
gtcattgcag GATTCATTCC CAGCCAGGTT CAAAGCTATC CACTTCATCC
ACCAGCCATG GTACTTCACC ACCACCTATA ATGTGGTCAA GCCCTTCTTG
AAGAACAAGC TGCTACAGAG Ggtgagtgtg tggctgcctc ccaacctttc
caaagaatga tcccaatttg ggtgaataat cctggacagc accacaagga
ggctgtagag agagacgtca cgtgacaaag atgggaggct gtgccttggt
ctcccctccc ccatccttgt gctcagttgc cctgctccct tgtctgcag
Exon 8
GTCTTTGTTC ACGGAGATGA CCTGGATGGC TTCTTCCAGG AGATTGATGA
GAACATCCTG CCTGCTGACT TTGGGGGTAC ACTGCCCAAG TACGACGGCA
AAGTTGTTGC TGAGCAGCTC TTTGGTCCCC GGGCTGAAGT TGAGAACACA
GCCTTATGAG GAGATGCCCT GCCATGGGAT CCGTATTAGC ATCCCAGTCC
TTCTTCAGCT TTCCTGAACC CAAGTCCGGG GACAGGAATG CCTGAAGGGA
CTGTTTTCCC AGACTTCCCT AGGTTTAGGA ACGCTCTAGT GGCACCTCAT
CCCAAACCGA CCTAGAGGGC AGTGAGGCAG GGGAATGCCT TTCATCGGGA
GATATGTGGG GCAGTTAAAT TCCCACCTGC CTCTGCAGTT ACAGGGCAGT
GACTCTGAGA GAAGTTGGAT TTCAAAGACT AATTAGGCTT CTCCGTGATT
TCCTTTGTAA TCGCTTTGAT ACTTAGGGCG TGGAAGGCAG AGTTAAAGGC
ATACTTGCAA TTCAGTAAAA AGAAGAAAGA GAGGGGCTTG AAGCTCTGAT
GATCTTAACA GCAGTAGGCA GATAAGGGGT TTCCCACCCA GGATTCCAGG
GAGGGGCCAG CAACCTTAGC ACAGTTGCTG ATGAATACTC TAGTAGCTAA
GAGCTCTGGA CTCTGTCCAG GGTGGAGGTC ATGTCCCTTC CTCTCCACCC
ATCTTGGATG TGATCCTTGG TGCTGGGGCC CCCATTCACT AGATGCTGTA
GGCTTTGTTG TATGGTTTTA CTTAAATGTA CCTTTGACAT CTTCCCCGAA
CACTCTCAGT TCTTATTAGG CAAATAAAAC CGTTCTTTGT CTTGTGTC
```

Figure 3. Mouse CRALBP Gene Sequence. The determined sequence (6855 nucleotides) of the 129/SvJ mouse CRALBP gene is shown with 5' flanking regions and introns in lower case letters and exons in upper case letters. Translated sequences are in green. Based on homology with the human CRALBP gene, the mouse transcription start site is numbered +1. The translation start-site (ATG) in exon 2, translation stop-site (TGA), polyadenylation signal (AATAAA) and terminator sequence (CTTGTGTC) are in red. The Genbank accession numbers for the sequences reported are: AF084638 (5' flanking regions and exons 1-3); AF084639 (exon 4); AF084640 (exon 5); AF084641 (exon 6); AF084642 (exons 7 and 8).

determined in both directions at least twice. Sequence results were interpreted and aligned using SEQUENCHER 3.0 software (Gene Codes, Ann Arbor, MI) and GeneWorks v 2.45N (Intelligenetics, Mountain View, CA) computer software. Sequence upstream of the presumptive translational start-site in the mouse CRALBP gene was inspected for consensus transcription factor binding sites using SIGNAL SCAN and Mat Inspector.

RESULTS

Copy Number and Transcript Size of the Mouse CRALBP Gene— Northern analysis reveals a predominant CRALBP mRNA species of about 2.1 kb in mouse eye RNA (Figure 1A) but not in mouse liver or kidney RNA (not shown). Similar size human CRALBP mRNAs are predicted (2046 bp and 2269 bp) from the human gene sequence [7]. Southern analysis of mouse 129/SvJ genomic DNA digested with different enzymes and hybridized with a 230 bp probe from the mouse CRALBP gene revealed a banding pattern totaling less than 20 kb/digest and consistent with a single copy gene (Figure 1B). Somatic cell hybridization previously localized the mouse CRALBP locus to a single site on mouse chromosome 7 [6].

Isolation of Mouse CRALBP Genomic Clones and cDNA Sequences— Using human and bovine CRALBP cDNA probes, a λ-phage clone designated λA2 was isolated from a mouse 129/Sv genomic DNA library and found to contain an ~13 kb insert that included the 5' end of the CRALBP gene through exon 6 (Figure 2). Three plasmid subclones were prepared from λA2 DNA and designated mp3.2SK, mp2.8SK and mp3.6CRII (Figure 2). The structures of mouse CRALBP exons 7 and 8 were obtained by direct sequence analysis of PCR products generated from mouse eye cDNA.

Characterization of the Mouse CRALBP Gene— The genomic organization of the mouse CRALBP gene is shown in Figure 2 and the DNA sequence of the gene in Figure 3. The sequence of the mouse CRALBP DNA was determined by automated analysis of genomic subclones and PCR products. About 5.3 kb of contiguous sequence was determined from overlapping subclones mp3.6CRII and mp2.8SK (putative gene positions -4300 to + 1000), including 4 kb of 5'-flanking sequence, exons 1-3 and introns 1-2. The sequences of exon 4 (205 bp) and exon 5 (179 bp) were determined from subclone mp3.2SK. The sequence of 1297 nucleotides of the mouse CRALBP cDNA was determined by direct analysis of PCR products generated from C57BL6 mouse eye cDNA. Mouse 129/SvJ genomic DNA was amplified with primers designed from mouse C57BL6 CRALBP cDNA, and direct sequence analysis of PCR products yielded the genomic sequences of exons 6-8, intron 7 (177 bp) and 692 bp of 3'-flanking sequence. The lengths of mouse CRALBP intron 3 (1.8 kb), intron 4 (1.5 kb), intron 5 (2.6 kb) and intron 6 (1.2 kb) were determined by PCR amplification using as a template either the 13 kb λ-phage clone or mouse 129/SvJ genomic DNA (Figure 4).

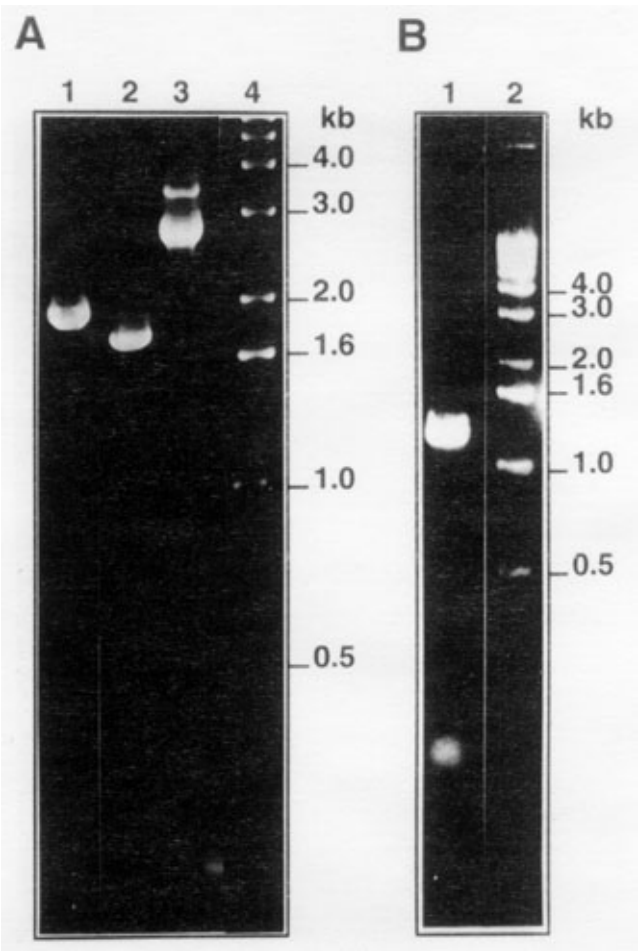

Figure 4. Mouse CRALBP Introns 3-6. The approximate size of mouse CRALBP introns 3-6 was determined by PCR amplification using primers designed to the flanking exons. (A) The ~13 kb mouse CRALBP phage clone as template: lane 1, intron 3 ~1.8 kb; lane 2, intron 4 ~1.5 kb; lane 3, intron 5 ~2.6 kb and lane 4, 1 kb DNA Ladder (Gibco, BRL). (B) 129/SvJ genomic DNA as template: lane 1, intron 6 ~1.2 kb; and lane 2, 1 kb DNA Ladder (Gibco, BRL). Exon sequence (200-300 nt) was subtracted from each of the observed intron PCR products.

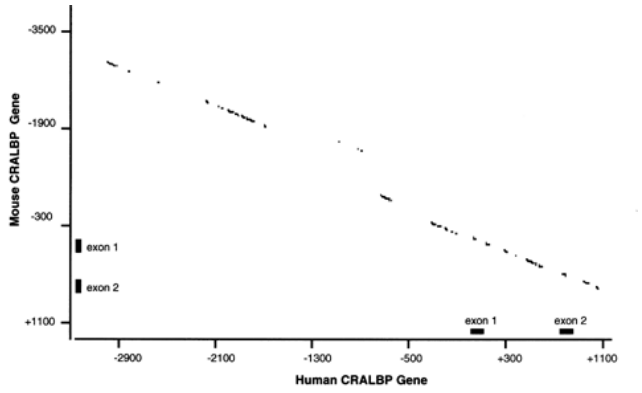

Figure 5. Mouse and Human CRALBP 5' Flanking Regions. Dot matrix comparison of human and mouse DNA sequences upstream of CRALBP exon 3. Gene positions are based on the human transcription start-site (+1 bp). Alignment was generated using GeneWorks program v2.5 (Intelligenetics) with a range of 16 nucleotides and stringency of 65% identity.

DISCUSSION

The structure of the mouse CRALBP gene was determined from a clone isolated from a 129/Sv mouse genomic library and from PCR amplification of mouse eye cDNA and mouse 129/SvJ genomic DNA. The mouse CRALBP gene is similar to the human gene in organization and size and contains eight exons and seven introns. In both genes, exon 1 is entirely untranslated, and the start (initiation) and stop (termination) codons are located in partially untranslated exons 2 and 8. A total of 6855 nucleotides of sequence was determined, including 3932 bases from the 5'-flanking region, 629 bases from the 3'-flanking region, all 8 exons and 3 introns. All mouse CRALBP exon/intron boundaries follow standard vertebrate GT/AG splicing convention [14]. Moreover, the 5' splice site of all introns conforms with the consensus 6 bp binding sequence of the spliceosome component U1 snRNP [15]. Exons range in size from 111 to 848 bases and introns vary from 178 to 2600 bases. Based on Southern blot analysis and localization to a single site on mouse chromosome 7 [6], the mouse CRALBP gene (*Rlbp1*) exists in single copy in the mouse genome. A search of the Mouse Genome Informatics (MGI) Resource at The Jackson Laboratory (Bar Harbor, ME) shows that no known mouse retinal diseases map to the CRALBP locus. The transcription start site of the mouse gene is predicted based on homology with the human CRALBP gene and is about 841 bases upstream of the ATG translation initiation signal. This putative transcription start site is consistent with the mouse CRALBP mRNA species detected by Northern analysis (~2.1 kb), a size within experimental error of the non-polyadenylated mRNA species (1868 bases) predicted from the mouse gene sequence. In the mouse gene, a consensus polyadenylation signal (AATAAA) and a terminator sequence (YGTGTTYY), separated by 12 bases, were found beginning 663 bases downstream from the translation stop (Figure 3). In the mouse CRALBP cDNA, a poly-A tail was found immediately downstream of the terminator sequence. Mouse and human CRALBP 3' untranslated regions are ~60% conserved. Unlike the gene encoding human RPE65, an RPE-specific protein also thought to be associated with vitamin A metabolism, the mouse and human CRALBP genes do not contain multiple ATTTA motifs implicated in transcript instability [16]. Overall, coding DNA in the mouse and human CRALBP genes exhibits ~87% sequence identity and non-coding DNA (exons 1-2) exhibits ~65% identity.

As an approach to detecting *cis*-elements that regulate CRALBP expression, evolutionary conserved sequences in the promoter region of the mouse and human genes have been

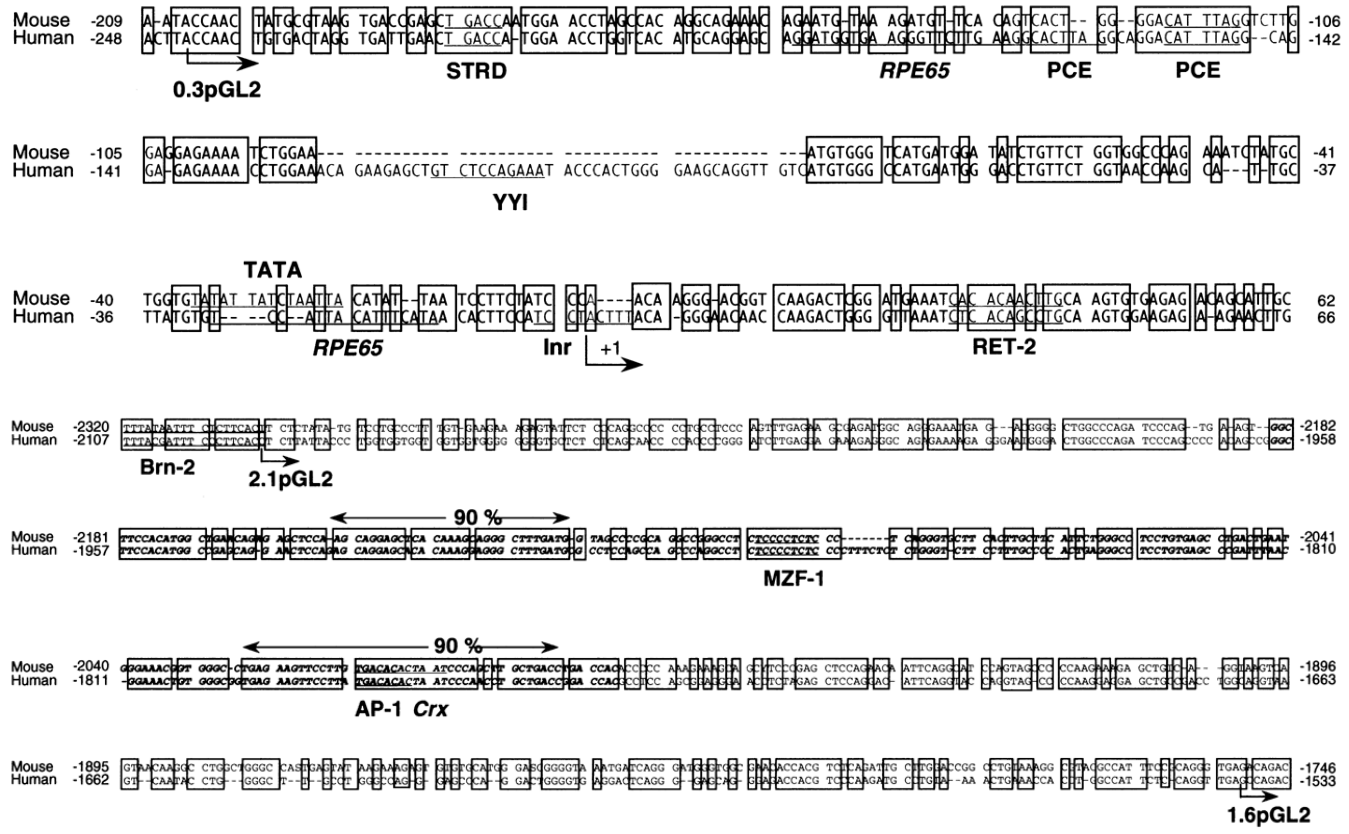

Figure 6. Proximal and Distal CRALBP Promoter Sequences. (A) Alignment of mouse and human CRALBP proximal promoter sequences. Conserved nucleotides are boxed and consensus regulatory elements are underlined. Represented by arrows are the transcription start-sites (+1) and the 5' end of the 0.3pGL2 human CRALBP reporter construct that directs RPE specific reporter expression in vitro [8]. The mouse gene transcription start-site is based on the human gene. (B) Alignment of mouse and human CRALBP distal promoter sequences. Conserved nucleotides are boxed, consensus regulatory elements are underlined and regions with >=90% nucleotide conservation are indicated by arrows. A previous study using human CRALBP reporter constructs 2.1 and 1.6 pGL2 (5' ends indicated) suggests the distal promoter region contains an enhancer element [8]. Sequence alignments were performed using GeneWorks 2.5 (Intelligenetics).

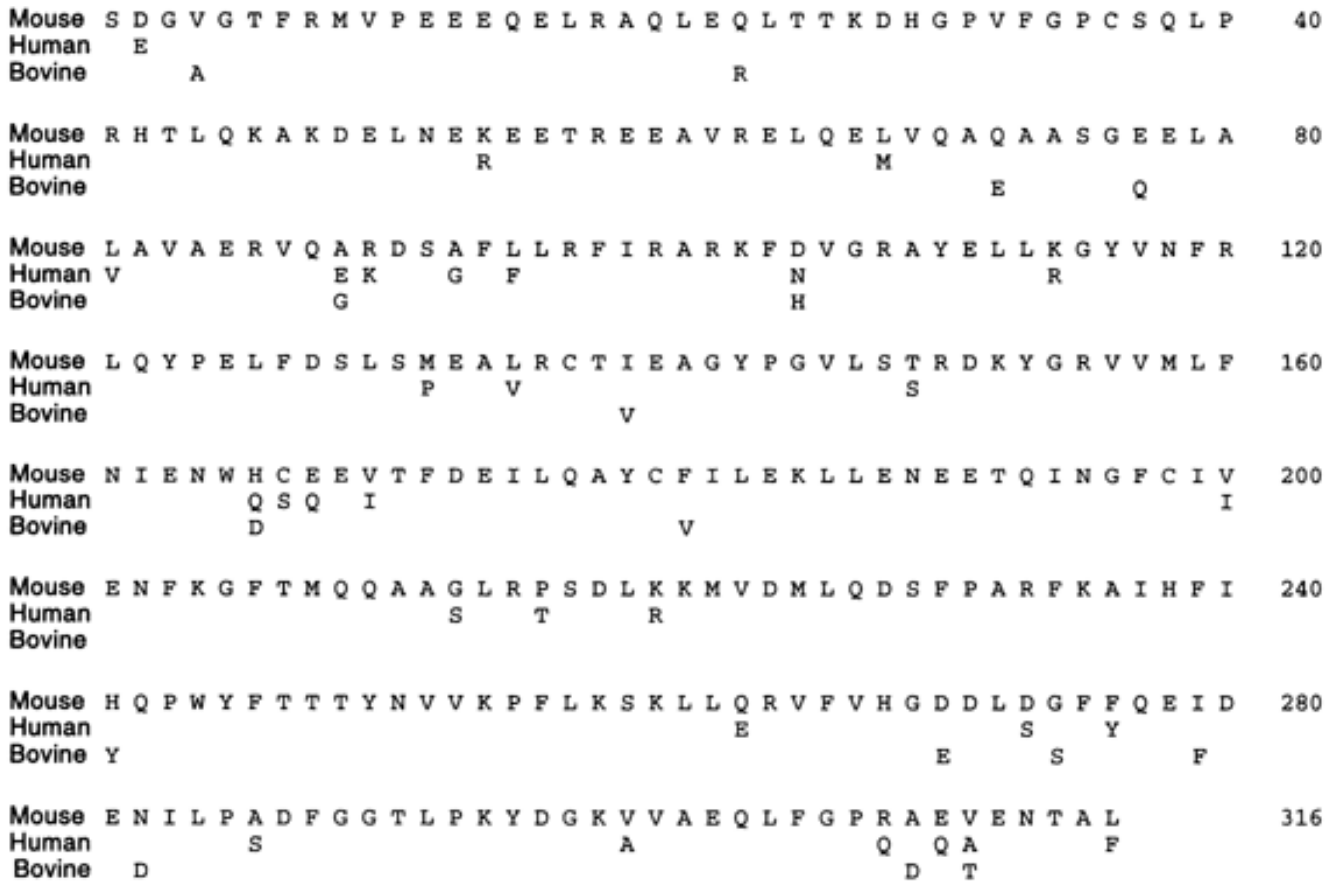

Figure 7. Comparison of CRALBP Protein Sequences. Alignment of the mouse, human and bovine CRALBP amino acid sequences is shown; residues not conserved with the mouse protein are indicated.

sought. Alignment of ~4 kb of mouse and human CRALBP gene sequences upstream of exon 3 reveals two homologous domains of ~1.4 kb and ~1.8 kb (approximate gene positions -2.9 kb to -1.5 kb and 0.7 kb to +1.1 kb, respectively), separated by the human-specific *Alu* repetitive sequence (Figure 5). Previously we reported that PCE1 sites in the human proximal promoter bind RPE nuclear factors that appear to interact with the distal promoter region -2089 to -1539 to drive high levels of CRALBP expression in the RPE [8]. The PCE1 consensus sequence CAATTAG (designated RCS1 in *Drosophila*) was originally identified in several mammalian photoreceptor genes and in *Drosophila* rhodopsin genes and proposed to direct photoreceptor cell-specific gene expression [17]. The two PCE1 between human *RLBP1* positions -165 and -140 exhibit significant homology with Ret-1 in the rat opsin gene, which has been reported to direct in vivo gene expression in both rod photoreceptors and brain [18]. One of the PCE1 sites is conserved in the mouse proximal promoter (Figure 6A), supporting the hypothesis that PCE1 influences CRALBP gene expression. Notably, two identical sequences (GCAGGA) flanking the PCE1 in human *RLBP1* and important for complex formation with human PCE1 binding proteins [8] are not conserved in the mouse gene. A consensus TATA box exists in the mouse gene [19] while the human proximal promoter

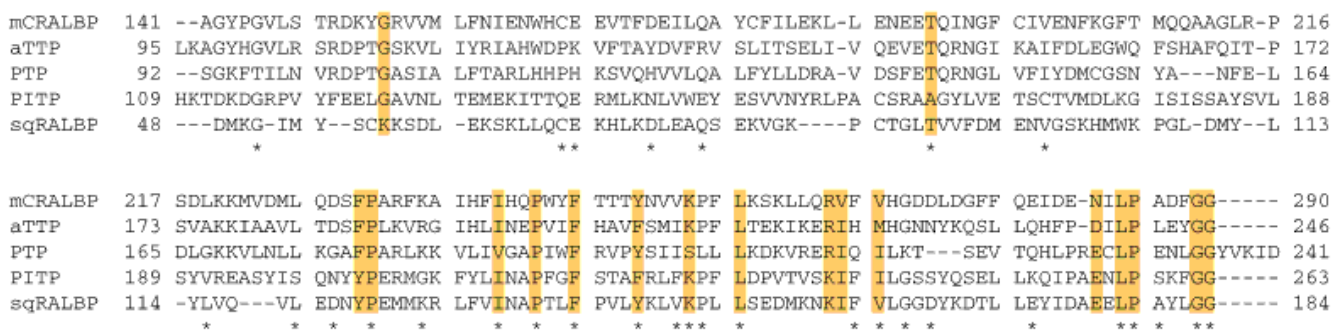

Figure 8. Structural Relatedness of CRALBP with other proteins. Homology is shown between mouse CRALBP, human α-tocopherol transfer protein (aTTP) [38], human protein tyrosine phosphatase (PTP) [39], yeast phosphatidylinositol-transfer protein (PITP) [40] and squid retinaldehyde-binding protein (sqRALBP) [33]. Numbers represent amino acid sequence positions; asterisks represent identical residues in mouse CRALBP and squid RALBP. Alignment was performed using GeneWorks program v2.5 (Intelligenetics).

contains an apparent Initiator (Inr) element, a 46 bp insert containing a Ying-Yang 1 (YYI) consensus site [20], and two RPE65 related sequences [21]. Other conserved proximal promoter consensus elements shown in Figure 6A include (i) RET-2, which in the rat opsin promoter specifically binds retinal nuclear factors [22]; and (ii) STRD, associated with high levels of arrestin expression in the retina and more broadly recognized by the vitamin D3, steroid, thyroid hormone and retinoic acid receptors [17,23]. The significance of these sites remains to be determined.

Conserved sequences in the mouse and human distal promoter region (Figure 6B) may be associated with putative enhancer activity in the human CRALBP gene between -2089 and -1539 [8]. Conserved distal promoter consensus sites include: (i) the Cone rod homeobox (Crx)-binding element that binds transcription factor Crx and transactivates several pineal/photoreceptor-specific genes [24,25,26]; (ii) activator protein 1 (AP-1) a common transcriptional activator [27]; (iii) Brn-2, associated with development and survival of the hypothalamus and pituitary [28,29]; (iv) MZF-1, important in tissue-specific expression of myeloid cells [30]; and (v) *Glass*-like (human CRALBP -2484 to -2465), implicated in photoreceptor cell development [31]. Further studies are required to determine the significance of these consensus sites and the most highly conserved sequence between *RLBP1* positions -1930 to -1900 bp and -1892 to -1852 bp.

The mouse, human and bovine CRALBP protein sequences each encode 316 amino acids and overall are about 87% identical (Figure 7). Important amino acid residues conserved in all three species include retinoid-binding pocket components Gln-210 and Lys-221 [32] and Arg-150, which if substituted with Gln (i. e., R150Q) can result in retinal degenerations associated with retinitis pigmentosa [5]. As previously described for bovine and human CRALBP, mouse CRALBP shares limited homology (Figure 8) with α-tocopherol transfer protein, protein tyrosine phosphatase and phosphatidyl inositol-transfer protein [7]. Contrary to an earlier report [33], CRALBP is also distantly related with squid retinaldehyde-binding protein as shown in Figure 8. Structure function studies with bovine and human recombinant CRALBP have characterized ligand interactions [3,32,34,35,36] and established in vitro evidence for a substrate carrier function in RPE [1,2,37].

This study represents initial work toward developing an in vivo model for studying CRALBP function and gene regulation in transgenic mice. The current results and clones have facilitated the construction of a mouse CRALBP gene targeting vector containing the neomycin resistance gene inserted into exon 3. Efforts to disrupt the mouse CRALBP gene by homologous recombination are in progress and homozygous knockout mice will be used for morphological, electrophysiological and retinoid metabolism studies. The current results also establish a foundation for identifying in vivo the minimal CRALBP promoter required for high level RPE expression in mice. Identification of the CRALBP minimal promoter will enhance gene therapy approaches for retinal diseases associated with CRALBP and other RPE expressed genes.

ACKNOWLEDGEMENTS

This study was supported in part by USPHS grants EY06603, EY02317, EY01730 and by Research to Prevent Blindness, Inc (RPB). JCS is a Senior Scientific Investigator of RPB. We gratefully acknowledge the technical assistance of S. Goldflam, P. S. Adams and A. Moquin and V. Oliver, and M. LaDuke in manuscript preparation. BK also thanks C. Morrissey and Drs. D. Eisinger, F. Martin and M. Tenniswood for useful discussions.

A preliminary report of this work was presented at the 70th Annual Meeting of the Association for Research in Vision and Ophthalmology, May 1998, Ft. Lauderdale, FL.

This work includes thesis research submitted by BNK in partial fulfillment of the requirements for the doctoral degree in the Cell and Molecular Biology Program jointly administered by the University College Dublin, Ireland and the Adirondack Biomedical Research Institute, Lake Placid, NY.

REFERENCES

1. Saari JC, Bredberg DL, Noy N. Control of substrate flow at a branch in the visual cycle. Biochemistry 1994; 33:3106-3112.
2. Winston A, Rando RR. Regulation of isomerohydrolase activity in the visual cycle. Biochemistry 1998; 37:2044-2050.
3. Saari JC, Bredberg L, Garwin GG. Identification of the endogenous retinoids associated with three cellular retinoid-binding proteins from bovine retina and retinal pigment epithelium. J Biol Chem 1982; 257:13329-13333.
4. Saari JC, Huang J, Possin DE, Fariss RN, Leonard J, Garwin GG, Crabb JW, Milam AH. Cellular retinaldehyde-binding protein is expressed by oligodendrocytes in the optic nerve and brain. Glia 1997; 21:259-268.
5. Maw MA, Kennedy B, Knight A, Bridges R, Roth KE, Mani EJ, Mukkadan JK, Nancarrow D, Crabb JW, Denton MJ. Mutation of the gene encoding cellular retinaldehyde-binding protein in autosomal recessive retinitis pigmentosa. Nat Genet 1997; 17:198-200.
6. Sparkes RS, Heinzmann C, Goldflam S, Kojis T, Saari JC, Mohandas T, Klisak I, Bateman JB, Crabb JW. Assignment of the gene (RLBP1) for cellular retinaldehyde-binding protein (CRALBP) to human chromosome 15q26 and mouse chromosome 7. Genomics 1992; 12:58-62.
7. Intres R, Goldflam S, Cook JR, Crabb JW. Molecular cloning and structural analysis of the human gene encoding cellular retinaldehyde-binding protein. J Biol Chem 1994; 269:25411-25418.
8. Kennedy BN, Goldflam S, Chang MA, Campochiaro P, Davis AA, Zack DJ, Crabb JW. Transcriptional regulation of cellular retinaldehyde-binding protein in the retinal pigment epithelium. A role for the photoreceptor consensus element. J Biol Chem 1998; 73:5591-5598.
9. Saari JC, Garwin GG, Van Hooser JP, Palczewski K. Reduction of all-trans-retinal limits regeneration of visual pigment in mice. Vision Res 1998; 38:1325-1333.
10. Crabb JW, Goldflam S, Harris SE, Saari JC. Cloning of the cDNAs encoding the cellular retinaldehyde-binding protein from bovine and human retina and comparison of the protein structures. J Biol Chem 1988; 263:18688-18692.
11. Don RH, Cox PT, Wainwright BJ, Mattick JS. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 1991; 19:4008.
12. Brown T. Preparation and analysis of DNA. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Struhl K,

editors. Current protocols in molecular biology. New York: John Wiley & Sons; 1993. p. 2.9.1-2.10.16.
13. Struhl K. Enzymatic manipulation of DNA and RNA. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Struhl K, editors. Current protocols in molecular biology. New York: John Wiley & Sons; 1995. p. 3.0.1-3.19.8.
14. Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA. Splicing of messenger RNA precursors. Annu Rev Biochem 1986; 55:1119-1150.
15. Staley JP, Guthrie C. Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell 1998; 92:315-326.
16. Nicoletti A, Wong DJ, Kawase K, Gibson LH, Yang-Feng TL, Richards JE, Thompson DA. Molecular characterization of the human gene encoding an abundant 61 kDa protein specific to the retinal pigment epithelium. Hum Mol Genet 1995; 4:641-649.
17. Kikuchi T, Raju K, Breitman ML, Shinohara T. The proximal promoter of the mouse arrestin gene directs gene expression in photoreceptor cells and contains an evolutionary conserved retinal factor-binding site. Mol Cell Biol 1993; 13:4400-4408.
18. Yu X, Leconte L, Martinez JA, Barnstable CJ. Ret 1, a cis-acting element of the rat opsin promoter, can direct gene expression in rod photoreceptors. J Neurochem 1996; 67:2494-2504.
19. Zenzie-Gregory B, O'Shea-Greenfield A, Smale ST. Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. J Biol Chem 1992; 267:2823-2830.
20. Pavletich NP, Pabo CO. Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. Science 1993; 261:1701-1707.
21. Nicoletti A, Kawase K, Thompson DA. Promoter analysis of RPE65, the gene encoding a 61 kDa retinal pigment epithelium-specific protein. Invest Ophthalmol Vis Sci 1998; 39:637-644.
22. Yu X, Chung M, Morabito MA, Barnstable CJ. Shared nuclear protein binding sites in the upstream region of the rat opsin gene. Biochem Biophys Res Commun 1993; 191:76-82.
23. Evans RM. The steroid and thyroid hormone receptor superfamily. Science 1988; 240:889-895.
24. Furukawa T, Morrow EM, Cepko CL. Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. Cell 1997; 91:531-541.
25. Chen S, Wang Q-L, Nie Z, Sun H, Lennon G, Copeland NG, Gilbert DJ, Jenkins NA, Zack DJ. Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. Neuron 1997; 19:1017-1030.
26. Li X, Chen S, Wang Q, Zack DJ, Snyder SH, Borjigin J. A pineal regulatory element (PIRE) mediates transactivation by the pineal/retina-specific transcription factor CRX. Proc Natl Acad Sci U S A 1998; 95:1876-1881.
27. Mitchell PJ, Tjian R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 1989; 245:371-378.
28. Nakai S, Kawano H, Yudate T, Nishi M, Kuno J, Nagata A, Jishage K, Hamada H, Fujii H, Kawamura K, Shiba K, Noda T. The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. Genes Dev 1995; 9:3109-3121.
29. Schonemann MD, Ryan AK, McEvilly RJ, O'Connell SM, Arias CA, Kalla KA, Li P, Sawchenko PE, Rosenfeld MG. Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. Genes Dev 1995; 9:3122-3135.
30. Hui P, Guo X, Bradford PG. Isolation and functional characterization of the human gene encoding the myeloid zinc finger protein MZF-1. Biochemistry 1995; 34:16493-16502.
31. Moses K, Rubin GM. Glass encodes a site-specific DNA-binding protein that is regulated in response to positional signals in the developing Drosophila eye. Genes Dev 1991; 5:583-593.
32. Crabb J, Nie Z, Chen Y, Hulmes JD, West KA, Kapron JT, Ruuska SE, Noy N, Saari JC. Cellular retinaldehyde-binding protein ligand interactions. gln-210 and lys-221 are in the retinoid binding pocket. J Biol Chem 1998; 273:20712-20720.
33. Ozaki K, Terakita A, Ozaki M, Hara R, Hara T, Hara-Nishimura I, Mori H, Nishimura M. Molecular characterization and functional expression of squid retinal-binding protein. A novel species of hydrophobic ligand-binding protein. J Biol Chem 1994; 269:3838-3845.
34. Saari JC, Bredberg DL. Photochemistry and stereoselectivity of cellular retinaldehyde-binding protein from bovine retina. J Biol Chem 1987; 262:7618-7622.
35. Luck LA, Barrows SA, Venters RA, Kapron JT, Roth KE, Barrows SA, Paradis SG, Crabb JW. NMR methods for analysis of CRALBP retinoid-binding. In: Marshak DR, editor. Techniques In Protein Chemistry VIII. San Diego: Academic Press; 1997. p. 439-448.
36. Crabb JW, Carlson A, Chen Y, Goldflam S, Intres R, West KA, Hulmes JD, Kapron JT, Luck LA, Horwitz J, Bok D. Structural and functional characterization of recombinant human cellular retinaldehyde-binding protein. Protein Sci 1998; 7:746-757.
37. Saari JC, Bredberg L. Enzymatic reduction of 11-cis-retinal bound to cellular retinal-binding protein. Biochim Biophys Acta 1982; 716:266-272.
38. Arita M, Sato Y, Miyata A, Tanabe T, Takahashi E, Kayden HJ, Arai H, Inoue K. Human alpha-tocopherol transfer protein: cDNA cloning, expression and chromosomal localization. Biochem J 1995; 306:437-443.
39. Gu M, Warshawsky I, Majerus PW. Cloning and expression of a cytosolic megakaryocyte protein-tyrosine-phosphatase with sequence homology to retinaldehyde-binding protein and yeast SEC14p. Proc Natl Acad Sci U S A 1992; 89:2980-2984.
40. Dundon W, Islam K. Nucleotide sequence of the gene coding for SEC14p in Candida (torulopsis) glabrata. Gene 1997; 193:115-118.