Identification of the RLBP1 Gene Promoter

John S. Vogel,1,2 Elizabeth C. Bullen,1 Crystal L. Teygong,1 and Eric W. Howard1

PURPOSE. Cellular retinaldehyde-binding protein (CRALBP), transcribed from the RLBP1 gene, is a 36-kDa water-soluble protein with 316 amino acids found in the retinal pigment epithelium (RPE) and in retinal Müller cells. It is thought to play a critical role in the visual cycle by functioning as an acceptor of 11-cis-retinol from the isomerohydrolase reaction. The goal here was to evaluate the functional promoter of this gene.

METHODS. 5′ RACE analysis, promoter-reporter assays, and semiquantitative PCR with exon-specific primers were performed using human-derived RPE cells (ARPE-19 and D407) in culture to evaluate the 5′ sequence flanking the RLBP1 gene. In addition, the murine, bovine, and porcine RLBP1 genes were evaluated in silico to identify likely proximal promoter/exon 1 sequences similar to the human gene.

RESULTS. 5′ RACE analysis revealed the presence of a previously undescribed exon in the RLBP1 gene. This was confirmed by analysis of the GenBank Human EST database, which revealed the presence of 18 sequences matching exon 1. Exon-specific PCR revealed that most CRALBP transcripts expressed in ARPE-19 cells contain both exon 1 and the final exon, suggesting that the primary promoter of CRALBP exists 5′ of the newly identified exon 1. Highly homologous sequences in the murine, bovine, and porcine genes were also identified. Finally, promoter-reporter constructs revealed a minimal sequence necessary for promoter function and indicated significantly greater promoter activity compared with previously described RLBP1 promoters.

CONCLUSIONS. The findings presented here suggest that CRALBP transcripts in RPE cells contain a noncoding exon in addition to a newly described promoter and, by definition, an additional intron. This finding sets the stage for a mechanistic understanding of the high degree of cell type–specific expression of RLBP1. (Invest Ophthalmol Vis Sci. 2007;48:3872–3877) DOI: 10.1167/iovs.06-1523

The retinal pigment epithelium (RPE) performs a number of critical functions within the retina, among them participating in the formation of the blood-retina barrier, the phagocytosis of distal portions of rod and cone outer segments and acting as an intermediary/transporter of ions/nutrients between the choroid and the neural retina.1 These cells are purported to have the ability to undergo phenotypic modulation, switching from an epithelial to a mesenchymal phenotype in diseases such as proliferative vitreoretinopathy.2–4 Molecular markers used to determine a switch from the proliferative to the differentiated phenotype include the expression of RPE-65 and cellular retinaldehyde-binding protein (CRALBP).5,6 These proteins are critically involved in retinoid (vitamin A) processing in the visual cycle, which includes the isomerization of all-trans-retinol to 11-cis-retinol. CRALBP, transcribed from the RLBP1 gene, is a 36-kDa water-soluble protein with 316 amino acids found in the RPE and in Müller cells of the retina. In the visual cycle, it is thought to function as an acceptor of 11-cis-retinol from the isomerohydrolase reaction and may facilitate its oxidation to 11-cis-retinal.7–9 Diseases associated with mutations in the RLBP1 gene include retinitis pigmentosa, fundus albipunctatus, Newfoundland rod/cone dystrophy, retinitis punctata albescens, and Bothnia dystrophy.10–14

The RLBP1 gene is located on chromosome 15q26 and was previously reported to contain 8 exons and 7 introns.15,16 A promoter for RLBP1 has been described that includes an upstream enhancer element responsible for a significant portion of promoter activity.16–18 This promoter was thought to initiate transcription 893 bases upstream of the translation start site based on RNase protection. However, little promoter activity was observed in constructs lacking the upstream enhancer. We provide bioinformatic and experimental evidence to suggest that, in fact, an additional exon exists upstream of the originally described exon 1 and that sequence elements 5′ of this exon, which were originally described as an enhancer, are actually the primary promoter for this gene.

MATERIALS AND METHODS

Cell Culture

ARPE-19 cells (passages 25–35; American Type Culture Collection [ATCC]; Manassas, VA) were cultured in 5% CO2 on 100-mm plastic dishes in DMEM/F12 (Gibco, Carlsbad, CA), 10% heat-inactivated fetal bovine serum (Life Technologies, Gaithersburg, MD), 100 U/mL penicillin G, 100 μg/mL streptomycin sulfate, and 0.25 g/mL amphotericin B (Life Technologies). Cultures were maintained at 37°C for up to 3 days. D407 cells (a kind gift from Richard Hunt, passages 68–74) were cultured in 5% CO2 on 100-mm plastic dishes in Dulbecco modified Eagle medium (DMEM; Gibco, Carlsbad, CA) with 5% fetal bovine serum plus 100 U/mL penicillin G, 100 μg/mL streptomycin sulfate, and 0.25 μg/mL amphotericin B (Life Technologies). For differentiation studies, the D407 cells were grown in serum-free medium supplemented with 0.2% lactalbumin hydrolysate (Life Technologies) for up to 3 days.

5′ RACE

Total RNA from differentiated ARPE-19 cells was analyzed with the use of a cDNA amplification kit (Smart RACE; BD Biosciences Clontech, Mountain View, CA) according to the manufacturer’s instructions. RNA was reverse transcribed using oligo(dT) primers. After PCR amplification with a universal primer from the 5′ kit (RACE; BD Biosciences Clontech) and a CRALBP-specific primer (5′-TGGCTATTAGGAAGACACAGAGT-3′) Integrated DNA Technologies, Coralville, IA) located in exon 3, the amplicons were cloned into pBluescript SK+ and screened for inserts using T3 and T7 primers. Clones were then PCR amplified with a nested 5′ primer from the 5′ kit (RACE; BD Biosciences Clontech) and

Investigative Ophthalmology & Visual Science, August 2007, Vol. 48, No. 8
Copyright © Association for Research in Vision and Ophthalmology

Downloaded From: http://iovs.arvojournals.org/pdfaccess.ashx?url=/data/journals/iovs/933443/ on 06/24/2017
a 3' primer specific to CRALBP (5'-TTGACCTGGTCCAAGTTGTC-3'). To determine whether these clones included exons 1, 2, and 3, they were next amplified with nested primers specific for either exon 1 (5'-GTCAGGTTGCTGGTTTGG-3') or exon 2 (5'-CTTGAGAC-CCAGGGTCACCTTGG-3') and exon 3 (5'-CTTGGAAGAAAGGAGGATCTGGT-3'). In addition, five of the RACE clones, representing several product sizes, were sequenced. Four of the five were of sufficient length to extend into the newly proposed exon 1 (see Fig. 2B, asterisks).

### Semi quantitative PCR

Total RNA from differentiated ARPE-19 cells was isolated according to the guanidine thiocyanate method, as previously described.10 Total RNA from D407 cells was isolated using a kit-based method (Purescript RNA isolation kit; Gentra Systems, Inc., Minneapolis, MN). In both cases, RNA was reverse transcribed with the use of reverse transcriptase (Superscript RT II RNase-H; Invitrogen, Carlsbad, CA) according to the supplier's protocol plus a 30-minute, 52°C incubation before the inactivation step. Reverse transcription was performed with oligo(dT) or random hexamers. DNAse-treated total RNA was used with the random hexamer–primed reverse transcription reaction. Genomic DNA from ARPE-19 cells was isolated using a kit-based purification system (Versagene DNA; Gentra Systems, Inc., Minneapolis, MN). DNA and total RNA were quantified by spectrophotometry. The following primers were used for PCR detection: CRALBP, 5'-CCCTCACTGTCGTTACCCAGGC-3' (upper) and 5'-TTACCCATCCCCCAACTTGAGA-3' (lower); RPE-65, 5'-GTTGGAGATATCCTGCCCTAC-3' (upper) and 5'-AGATGATGATCCTGGGTTTGGAA-3' (lower); smooth muscle α-actin, 5'-AACACATAGGTAAAGGACTAGCAG-3' (upper) and 5'-GGAAGGAGCCCTATGTCAACTG-3' (lower); and GAPDH, 5'-CCATCCTGCCCACCAAG-3' (upper) and 5'-CAGGAATGCCCTTGAAG-3' (lower). The following primers were used for evaluation of CRALBP mRNA species after random–primed reverse transcription: primer 1, 5'-GTACACGGTAGACACTTCTCCTC-3'; primer 2, 5'--CTTGACGGACTCTCAGGTTTG-3'; primer 3, 5'-CTTGGAAGAAAGGAGGATCTGGT-3'. Primers 1 and 3 amplified a 238-bp product spanning exon 1 to exon 3. Primers 2 and 3 amplified a 146-bp product spanning exon 2 to exon 3. All PCR products were analyzed on agarose gels (GenePure LE; ISC BioExpress, Kaysville, UT) and were stained with ethidium bromide for visualization on a UV transilluminator (Chemilumager; Alpha Innotech Corp., San Leandro, CA).

### Luciferase Reporter Analysis

Promoter fragments were amplified from genomic or plasmid DNA using primers as follows (+1 represents the translation initiation site in exon 3): -1519/807, 5'-TCCCTCAAGCTTCTTTTG-3' (upper), 5'-CTTGAAACCCAGGTCAACCTTG-3' (lower); +3445/807, 5'-CAGTTAGGCCGCTTATGATGTG-3' (upper), lower is the same as for -1519/807; -4686/2568, 5'-AGCAACATCTCAAGTGG-3'; -3157/2568, 5'-CTTGCAAGCCCTCTTTTG-3'; -2568/807, 5'-AGCAACATCTCAAGTGG-3'; -3157/2568, 5'-AGTAATCGCTCACT-3' (lower); and -3157/2568, 5'-AGGAATCCAAAATGCGGAAATG-3' (upper), lower is the same as for -4686/2568. Constructs were ligated into pGL3 Basic (Promega, Madison, WI) and were cotransfected into D407 cells with a CMV promoter-driven β-galactosidase reporter gene vector with the use of a transfection reagent (Lipofectamine; Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions. Briefly, 0.6 μg each reporter construct cDNA and 0.2 μg β-galactosidase vector were transfected in quadruplicate. After 12 hours, the medium was changed to serum-free medium supplemented with lactalbumin hydrolysate. At 3 days, cells were lysed in luciferase reporter buffer. Luciferase luminescence was analyzed on a luminometer (TD-20/20; Turner Designs, Inc., Sunnyvale, CA). To correct for transfection efficiency, β-galactosidase activity was normalized using chlorophenol red-β-D-galactopyranoside (CPRG; BioChemi Imaging System; UVP BioImaging Systems, Upland, CA).

### RESULTS

#### 5' RACE Analysis of CRALBP Transcripts

Along with pigmentation genes and other members of the vision cycle, CRALBP is a differentiation marker for RPE cells.5,6 We evaluated CRALBP expression by cultured RPE cell models and found that, in ARPE-19 cells cultured in serum–containing medium, expression increases after day 9 and continues increasing through day 25 (Fig. 1A). RPE65 upregulation was significant by day 15. These changes in gene expression correlate with the cells attaining a cobblestone-like appearance and are similar to what has been previously described in long-term ARPE-19 cultures.9 In contrast, D407 cells, which are a spontaneously transformed human RPE cell line, rapidly upregulated the expression of RPE-65 and CRALBP within 2 to 3 days after they were switched to serum-free conditions (Fig. 1B).

The current GenBank entry for the human RLBP1 gene (NC_000015.8:87505926.87554102, complement) and the reference mRNA (NM_000526) suggest that this gene has an additional exon not previously reported. Based on the GenBank data, we have depicted the structure of the RLBP1 gene in Figure 2A, including two noncoding exons and nine total exons. Figure 2B illustrates the human gene aligned with the putative first exons of the murine, bovine, and swine RLBP1 genes, along with the exon 1–intron 1 boundaries. It also depicts what we now term exon 2 in mouse and human genes. Based on these alignments, we predict that, like the human gene, the murine, bovine, and swine RLBP1 genes initiate transcription at the 5' boundary of exon 1, as depicted
in Figure 2B. This is in contrast to what is predicted by GenBank for these three species. The existence of exon 1 and its 5’ flanking sequence is not included in the curated GenBank entries.

To confirm experimentally the presence of the predicted new exon 1 and to identify the functional RLBP1 promoter, we mapped the transcription start site of the human RLBP1 gene using 5’ RACE with ARPE-19 cells as a source of RNA. Several 5’ RACE products were cloned and sequenced and were found to initiate significantly 5’ of the previously reported transcription start site. This site is depicted in Figure 2B and falls just 5’ of what we now call exon 2. The 5’ ends of several of the
RACE products were found to localize to regions within exon 1 (Fig. 2B, asterisks), and the RACE clones themselves, which were amplified using a primer that anneals to sequence homologous to exon 3, showed consistent boundaries between exons 1 and 2, revealing the existence of novel exon 1 and intron 1 (Fig. 2B). Figure 2B also reveals the high degree of sequence homology in the promoter region and in exon 1 among the four species. To further confirm the identity of the proposed exon 1, we scanned GenBank for human expressed sequence tag (EST) clones that might contain sequence identical with this exon and found 18 unique EST clones that fulfilled this criterion (Fig. 3). Similar analysis of murine and bovine ESTs identified 9 clones for each species and two EST clones from swine (not shown), each containing sequence homologous to exon 1 (Fig. 3). Together, these results argue for the existence of an exon in this gene proximal to previously described exons.16

Most CRALBP Messages Contain Exon 1 and Exon 9

We next sought to determine whether CRALBP mRNA species containing the newly identified exon 1 represent a small fraction of total CRALBP transcripts or if they, in fact, represent a significant proportion of transcripts. To do this, we evaluated 23 bacterial colonies, based on the diagram in Figure 4A, containing CRALBP RACE products for the presence of exons 1 and 3. Of the 23 clones that contained exons 2 and 3 based on amplification with primers that span these two exons, 22 also contained exon 1 based on PCR amplification using primers that span exon 1 and exon 3 (Fig. 4B). Next, we developed PCR primers that would specifically amplify within exon 1 or the 3′ untranslated region of CRALBP and used these to determine whether the proportion of messages containing exon 1 at least equaled those containing the final exon. As shown in Figure 4C, both primer sets were similar in efficiency when used to amplify genomic DNA. Amplifying cDNA with these primers demonstrated that the proportion of exon 1–containing messages was similar to that of exon 9–containing mRNAs. Together, these results suggest that human RLBP1 gene transcription initiates at a site significantly upstream of the previously described promoter, resulting in a primary transcript containing an exon and an intron additional to those previously identified.
Sequence Flanking Exon 1 Is a Functional RLBP1 Promoter

Based on these findings, we evaluated the ability of genomic sequence flanking the putative exon 1 to support transcription in RPE cells. Previous studies located the proximal RLBP1 promoter to sequence flanking what we refer to as exon 2, and an enhancer located within the region that includes exon 1.16,17 A 2659-bp reporter construct spanning part of exon 2 through sequence 5′ of exon 1 supported transcription, as illustrated in Figure 5C (constructs -1519/-807 and -3445/-807). As previously reported, eliminating the sequence upstream of intron 1 provided only minimal promoter activity.17 We interpret this to mean that the bona fide transcription initiation sites are located in the region labeled promoter/exon 1 in Figure 5A. To further substantiate this, we assembled reporter constructs that excluded all intron 1 and exon 2 and that included exon 1 and the sequence adjacent to it (Fig. 5C). Three different constructs with progressively shorter promoter fragments yielded robust transcriptional activity (Fig. 5C; -4686/-2568, -3843/-2568, and -3157/-2568). Next, we deleted a small sequence at the 3′ end of exon 1 (between -2687 and -2568) and found that this eliminated promoter activity (Fig. 5C; -4686/-2687, -3843/-2687, and -3157/-2687). Further evaluation of the sequence included in this region revealed the presence of a number of putative transcription factor-binding sites, particularly a putative Sp1 site (Fig. 5B). Progressive deletions of the 3′ end of the -3157/-2568 promoter fragment revealed the critical importance of sequence including the Sp1 site; excluding this region resulted in a significant decrease in promoter activity (Fig. 5D).

DISCUSSION

The RPE consists of highly differentiated epithelial cells that express proteins not found in most other cell types.1 CRALBP is one such marker, and its expression can be induced in ARPE-19 and D407 cell lines under appropriate conditions. Early studies of this gene suggested that the transcription initiation site was located -893 bases from the translation initiation site.16 We present evidence that, in fact, an additional exon not previously identified as such is transcribed from a novel promoter located in the region 5′ to -2558 from the translation start site. This promoter supports expression in a human RPE cell line with a differentiated phenotype in the absence of previously identified elements contained within intron 1. Evidence for the proposed sequence serving as the primary promoter for CRALBP expression includes RACE analysis, the existence of ESTs from multiple species that include the newly defined exon 1, functional analysis of promoter activity, and PCR analysis demonstrating that most CRALBP mRNAs contain exon 1.

Previous studies used primer extension and RNase protection assays to locate the RLBP1 promoter adjacent to what we have identified as exon 2.16 It was suggested that the human RLBP1 promoter contains a powerful upstream enhancer element located -922 bases from the translation initiation site.17

![Figure 5](http://iovs.arvojournals.org/pdfaccess.ashx?url=/data/journals/iovs/933443/)
Reporter constructs used to demonstrate this were similar to the -5445/-807 construct illustrated in Figure 5C, which includes exon 1, intron 1, and part of exon 2. As previously demonstrated, truncating this fragment leads to a profound loss of promoter activity.\textsuperscript{17} Based on our findings that the sequence upstream of exon 1 can itself act as a promoter, we interpret the previous findings as support for the existence of primary transcription initiation sites within the previously suggested enhancer region. This idea is strengthened by our findings that reporter constructs that exclude elements 3’ of exon 1 result in no promoter function above baseline. Thus, previous studies correctly identified the importance of this region for supporting transcriptional activity in RPE cells.\textsuperscript{13} It appears, however, that the reason for this is its importance as a transcription initiation site, not as an enhancer.

Additional support for the existence of the new exon 1 comes from conserved 5’ and 3’ splice sites that conform to consensus splice sequences. The splice sites revealed by our RACE studies and by EST database analysis of the bovine, swine, murine, and human genes include the completely conserved sequence, CAGGTGAGT, which is an optimal 5’ splice site, and a polyypyrimidine tract followed by an AG sequence, an optimal 3’ splice acceptor site.\textsuperscript{20} Together with promoter analysis, the presence of consensus splice sequences argues for the existence of the newly defined exon 1 and intron 1. CRALBP mRNA species were originally reported to measure approximately 1.4 and 1.6 kb, as determined by Northern blot analysis of human retinal RNA, and it was speculated that this was the result of multiple polyadenylation sites.\textsuperscript{16} In fact, the 1.6-kb species was predominant and corresponded in size to what we think is the primary mRNA species that includes exon 1.

The variability of apparent transcription start sites was based not only on our RACE analysis but also on EST clone sequences. Large-scale mapping of full-length mRNAs has revealed significant diversity in transcription initiation, particularly in TATA-less promoters.\textsuperscript{21} Indeed, our findings do not necessarily rule out the existence of other promoters supporting CRALBP expression. The possibility exists that the previously identified RLBP1 proximal promoter, which we think is actually part of intron 1, serves as an alternative promoter. Our studies making use of transient transfection analysis suggest that, if this is the case, the promoter 5’ of exon 2 provides only minimal activity compared with the promoter flanking exon 1. In addition, the bovine and swine genes lack the exon that flanks this putative promoter, and we could find no ESTs in the database representing RLBP1 that contain sequence homology with any part of bovine intron 1. Thus, murine and human RLBP1 genes have 9 identified exons; the translation initiation sites occur in exon 3. The bovine gene, which has 8 exons including the newly assigned exon 1, initiates the translation of CRALBP from exon 2 and lacks the intermediate noncoding exon that murine and human genes have. Sequence alignment of bovine, swine, murine, and human RLBP1 gene sequences, as shown in Figure 2, demonstrates significant homology in the promoter region flanking exon 1, including several putative transcription factor-binding sites that are highly conserved and several likely Sp1 sites. There is significantly less interspecies homology flanking human exon 2. The putative Sp1 site, located just 5’ of our RACE clones, is typical of many TATA-less promoters. Its striking conservation in the four species compared in Figure 2 further argues for its importance in CRALBP expression. This is further supported by the observation that the exclusion of this site from promoter constructs results in markedly decreased promoter activity in transiently transfected D407 cells (Fig. 5D). Further analysis will be required to evaluate the identities of other important transcription factor-binding sites and to identify elements involved in tissue-specific expression of this gene.