Cell, Vol. 34, 807–814, October 1983, Copyright © 1983 by MIT 0092-8674/83/100807-08 $02.00/0

Isolation, Sequence Analysis, and Intron–Exon Arrangement of the Gene Encoding Bovine Rhodopsin

Jeremy Nathans and David S. Hogness*
Department of Biochemistry
Stanford University School of Medicine
Stanford, California 94305

Summary

We have isolated cDNA clones generated from the mRNA encoding the opsin apoprotein of bovine rhodopsin and used these cDNAs to isolate genomic DNA clones containing the complete opsin gene. Nucleotide sequence analysis of the cloned DNAs has yielded a complete amino acid sequence for bovine rhodopsin and provided an intron–exon map of its gene. The mRNA homologous sequences in the 6.4 kb gene consist of a 96 bp 5′ untranslated region, a 1044 bp coding region, and a surprisingly long ~1400 bp 3′ untranslated region, and are divided into five exons by four introns that interrupt the coding region. Secondary structure analysis predicts that the bovine rhodopsin chain, like that of bacteriorhodopsin, contains seven transmembrane segments. Interestingly, three of the four introns are immediately distal to the codons for three of these segments, and one of these introns marks the boundary between the C-terminal domain and a transmembrane domain.

Introduction

Rhodopsin is the light-absorbing pigment in the retina responsible for visual excitation. It consists of an apoprotein, opsin, covalently linked to a conjugated polyene chromophore, 11-*cis* retinal. Rhodopsin is an integral membrane protein that lies embedded in the disc membranes of the photoreceptor outer segment. The chromophore in all visual pigments that have been analyzed is either 11-*cis* retinal (or 11-*cis* dehydroretinal). Given that different pigments employing the same chromophore exhibit different absorption spectra, it is generally supposed that this variation in absorbance results from differences in the structure of the opsin apoprotein (Wald, 1968). In particular, the trichromatic color vision of humans and other primates is imagined to require three different opsins, one for each of the three kinds of cones (blue, green, and red sensitive), each different from the rod opsin of rhodopsin, and all belonging to the 11-*cis* retinal family. In humans the existence of different opsins is further suggested by the well known genetic variation in color vision. Inherited color vision defects may decrease, shift, or abolish the spectral sensitivity of one of the cone mechanisms, and in general, alleles at each locus perturb only one of the three cone mechanisms (Boynton, 1979).

We want to test the hypothesis that different opsins mediate human color vision. Given the relative paucity of the cone visual pigments, a test by direct isolation of their opsins is difficult. Consequently, we have opted to test the hypothesis by isolation of the genes that encode these opsins—from both wild type and mutant. Analysis of the structure of these genes should allow detection of differences in the primary structure of the respective opsin proteins. Furthermore, it may be possible to synthesize these proteins by expression of the cloned genes in cell environments more favorable to their isolation, and thereby to test the effects of differences in primary structure on the spectral properties of these proteins when they are combined with 11-*cis* retinal.

Our strategy assumes that the different opsin genes possess sufficient sequence homology within and between mammalian species that a molecular clone of one gene can be used to isolate the others. More specifically, our experimental plan is first to isolate the gene encoding the best studied of these opsins, namely that in bovine rhodopsin; second, to use the coding sequences in that gene as a hybridization probe to identify and isolate at least one of the human opsin genes from a library of human genomic DNA clones; and third, similarly to use the first human gene isolate to identify and isolate the other members of the postulated opsin multigene family in humans.

The first step in the plan is reported in this paper. That step consists of the isolation and identification of bovine opsin cDNA clones, followed by the use of these clones to isolate cloned genomic DNA segments containing the complete bovine opsin gene. Nucleotide sequence analysis of the genomic and cDNA clones has provided both a complete amino acid sequence for bovine rhodopsin and an intron–exon map of its gene.

Results

Isolation of Opsin cDNA Clones

Libraries of cDNA clones were constructed by inserting into the λgt10 phage vector duplex cDNA segments prepared from polysomal poly(A)$^+$ RNA isolated from bovine retinas (Experimental Procedures). These libraries were successfully screened for opsin cDNA clones by three methods. A detailed description of one of these methods is given below, followed by summary descriptions of the other two.

A probe specific for opsin cDNA was prepared by reverse transcriptase extension from a synthetic oligodeoxynucleotide primer annealed to bovine opsin mRNA at the position encoding the five amino acids located −38 to −42 residues from the carboxy terminus of bovine opsin. Figure 1 shows this amino acid sequence, the 12 possible mRNA sequences that could encode it, and the 15-mer primer that we synthesized and used to prime reverse transcription. This primer was selected over the other 11 possibilities because Satterthwait, Oprian, and Khorana (personal communication) and Baehr and Applebury (1983) had observed priming of bovine opsin mRNA with

* To whom reprint requests should be addressed.

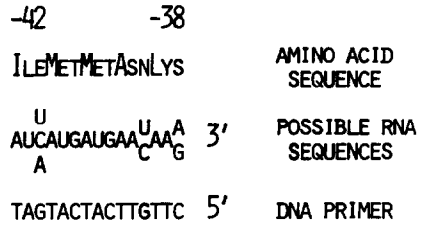

Figure 1. Sequence of the Synthetic DNA Primer

The synthetic DNA primer is compared with the possible mRNA sequences encoding the five amino acids located −38 to −42 residues from the carboxy terminus of bovine opsin (Hargrave et al., 1982). The positions are numbered from the carboxy terminus because the complete sequence was unknown at the time of these experiments.

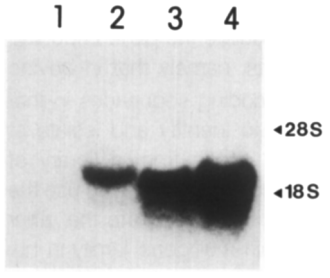

Figure 2. The DNA Primer Hybridizes to a Single Size Class of Polysomal Poly(A)$^+$ RNA

The four electrophoretic gel lanes contain (1) 40 μg total RNA from D. melanogaster embryos, a control; (2) 44 μg total RNA from bovine retinal polysomes; (3) 4 μg and (4) 8 μg poly(A)$^+$ RNA from bovine retinal polysomes. Conditions of hybridization with radioactive 15-mer are described in Experimental Procedures. The arrowheads indicate the mobilities of the bovine 18S and 28S rRNAs. The mobility of the RNA yielding the hybridization band common to lanes 2–4 is the same as that of authentic bovine opsin mRNA, which was assayed by hybridization with the cloned opsin cDNA segment, bd8, and whose 2.65 length was determined by comparison of its mobility with that of pBR322 length standards (Sutcliffe, 1978; McMaster and Carmichael, 1977). Exposure: 12 hr.

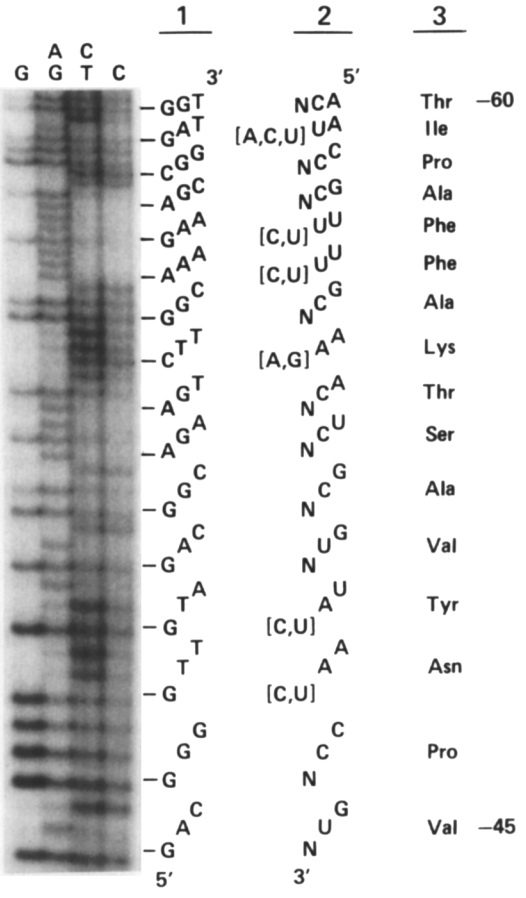

Figure 3. Comparison of the Nucleotide Sequence of the 15-Mer Primed cDNA Extension Product with the Possible Opsin mRNA Sequences

Column 1 shows the cDNA sequence derived from the autoradiograph of the sequencing gel. In this set of chemical cleavages, C and T are sometimes difficult to distinguish because of modification by residual hydrazine; in such cases, the choice between them was made on the basis of complementarity to the base in the mRNA sequence shown in column 2. The mRNA sequence derives from the opsin amino acid sequence for residues −45 to −60 from the carboxy terminus (Hargrave et al., 1982) shown in column 3. Exposure: 4 days.

a tetradecamer (5′-TGCTTGTTCATCAT-3′) a dodecamer (5′-GTTCATCATGAT-3′), respectively, that it overlaps.

Figures 2 and 3 demonstrate that this 15-mer hybridizes specifically to the opsin mRNA in poly(A)$^+$ RNA populations from retinal polysomes. Figure 2 shows that when such RNA populations are fractionated according to size by gel electrophoresis, blot transferred to diazatized paper (Alwine et al., 1977), and hybridized with the 5′-end ^{32}P-labeled 15-mer, almost all of the hybridization is restricted to a single RNA size class of 2.65 kb. When this same 5′-end-labeled 15-mer was hybridized to the poly(A)$^+$ RNA and used as a primer for cDNA synthesis by reverse transcriptase, the nucleotide sequence of the cDNA extension product is that expected for synthesis primed on the opsin mRNA, as is shown in Figure 3.

Nitrocellulose filter replicas (Benton and Davis, 1977) of the cDNA library were probed with 15-mer primed cDNA prepared as in Figure 3 except that its radioactivity derived from α-^{32}P-dCTP incorporated during the reverse transcription. Following autoradiography, the cDNA probe was eluted from the filters and a control cDNA probe, synthesized in a similar reaction without addition of the 15-mer, was hybridized to the filters (Figure 4). Using a 15-mer primed extension product as the probe, rather than the 15-mer alone, allowed the filters to be hybridized and washed under standard conditions (Experimental Procedures). Approximately 5% of the recombinants hybridized to both cDNA probes, and approximately 0.2% of the recombinants hybridized only to the 15-mer primed cDNA. Eight clones that hybridized only to the probe primed with the 15-mer were examined. Conceptual translation of the nucleotide sequence in two of the cloned cDNA segments (bd8 and bd20, Figures 5C and 6) proved that each was copied from opsin mRNA (see below), and cross-hybridization of the remaining six clones with the bd8 segment demonstrated that all are opsin cDNA clones (data not shown).

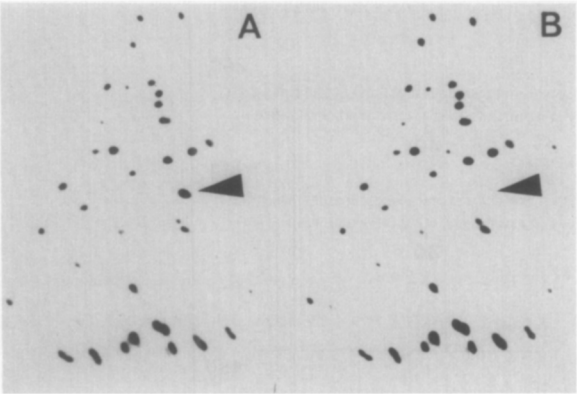

Figure 4. Screening the cDNA Library for Opsin cDNA Clones

A nitrocellulose filter replica of an agar plate containing 800 recombinant phage plaques from the library was first hybridized (Benton and Davis, 1977) with ^{32}P-labeled 15-mer primed cDNA, prepared as described in Experimental Procedures, to yield autoradiograph (A). The radioactive cDNA was removed by heating at 80°C for 15 min in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and the filter then hybridized with a control probe (see text), yielding autoradiograph (B). The arrowheads indicate a plaque that hybridized specifically with the 15-mer primed cDNA. Exposure: 15 hr.

Each of the other two methods for screening the cDNA libraries for opsin clones consisted in the differential hybridization of a pair of duplicate plaque filters with two similar ^{32}P-labeled cDNA probes that differ by the presence of opsin mRNA sequences in the first but not in the second. In one of these methods, the first cDNA probe was copied from retinal poly(A)$^+$ RNA, while the second derived from brain poly(A)$^+$ RNA. In the other method, the first and second cDNA probes derived from the poly(A)$^+$ RNA of retinal polysomes that did and did not, respectively, adhere to anti-rhodopsin IgG (Gough and Adams, 1978). In both experiments, those plaques that hybridized most strongly with the first probe and hybridized either weakly (in the second experiment) or not at all (in the first experiment) with the second probe were shown to be opsin clones by cross-hybridization with the bd8 opsin cDNA segment (data not shown).

Isolation of Opsin Genomic Clones

When Eco RI and Hind III digests of total bovine retinal DNA were fractionated by gel electrophoresis, blot transferred to nitrocellulose paper, and hybridized with ^{32}P-labeled (Rigby et al., 1977) bd8 cDNA under standard conditions (Experimental Procedures), hybridization to fragments of 18 kb (Eco RI) and 12 kb (Hind III) were observed (data not shown). Both fragments were isolated as recombinant λ phages by screening libraries of cloned Eco RI and Hind III genomic fragments with a ^{32}P-labeled cloned opsin cDNA probe (Experimental Procedures). Figure 5A shows the overlapping genomic arrangement of the cloned 18 kb Eco RI (Bd1) and 12 kb Hind III (Bd7) fragments. Although the above results suggest the possibility that the bovine genome contains only one gene coding for opsin proteins, we cannoit draw that conclusion because we have not attempted the identification of weakly homologous genomic fragments by cDNA hybridization at low stringency.

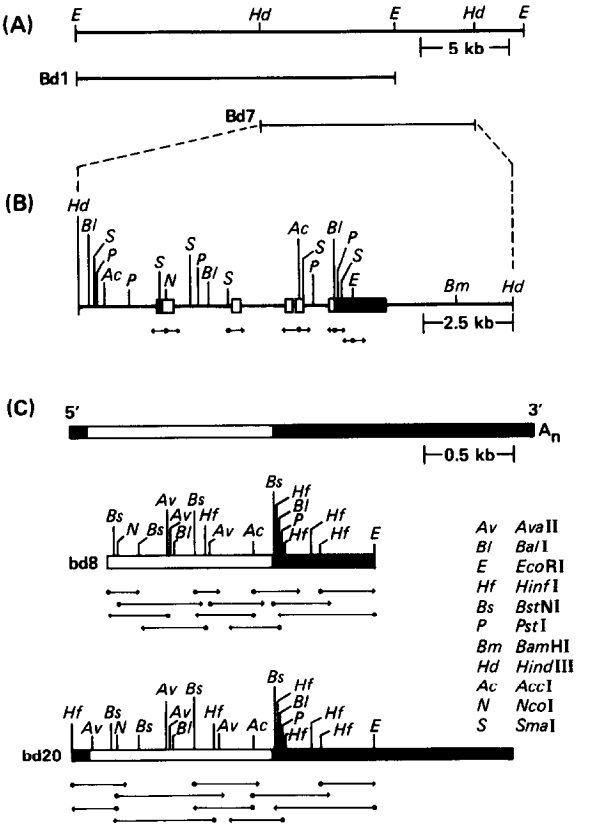

Figure 5. Restriction Maps and Sequencing Strategies

(A) The overlapping cloned 18 kb Eco RI (Bd1) and 12 kb Hind III (Bd7) segments are shown below a restriction map for the chromosomal DNA from which they derive. (B) The open and closed blocks in this expanded map of Bd7 represent, respectively, coding and untranslated regions within the five exons that are connected by horizontal lines representing four introns with the following lengths: intron 1 (leftmost), 1.65 kb; intron 2, 1.20 kb; intron 3, 0.117 kb; intron 4, 0.85 kb; see Figure 6 for exon lengths. A detailed restriction map is given for the region to the left of the Eco RI site, and below that is the sequencing strategy (Maxam and Gilbert, 1980), where each arrow indicates the determined sequence of the strand whose 3′-end-labeled site is marked by the dot at the tail of the arrow. (C) The mRNA structure is shown above the restriction maps and sequencing strategies upstream of the Eco RI site in the bd8 and bd20 cDNA segments. Symbols are the same as in (B).

Sequence Analysis of the Bovine Opsin Gene and Its mRNA

Figures 5B and 5C show our strategy for determining the nucleotide sequences of the indicated regions in the cloned genomic (Bd7) and cDNA (bd8 and bd20) segments. Figure 6 summarizes the results of these sequence determinations and of experiments mapping the mRNA start site in the opsin gene. This mapping was accomplished by both S1 nuclease protection and cDNA extension of genomic fragment strands which were 5′-end labeled at, and extend upstream from, the unique Nco I site (Figure 5B), and which were annealed to opsin mRNA

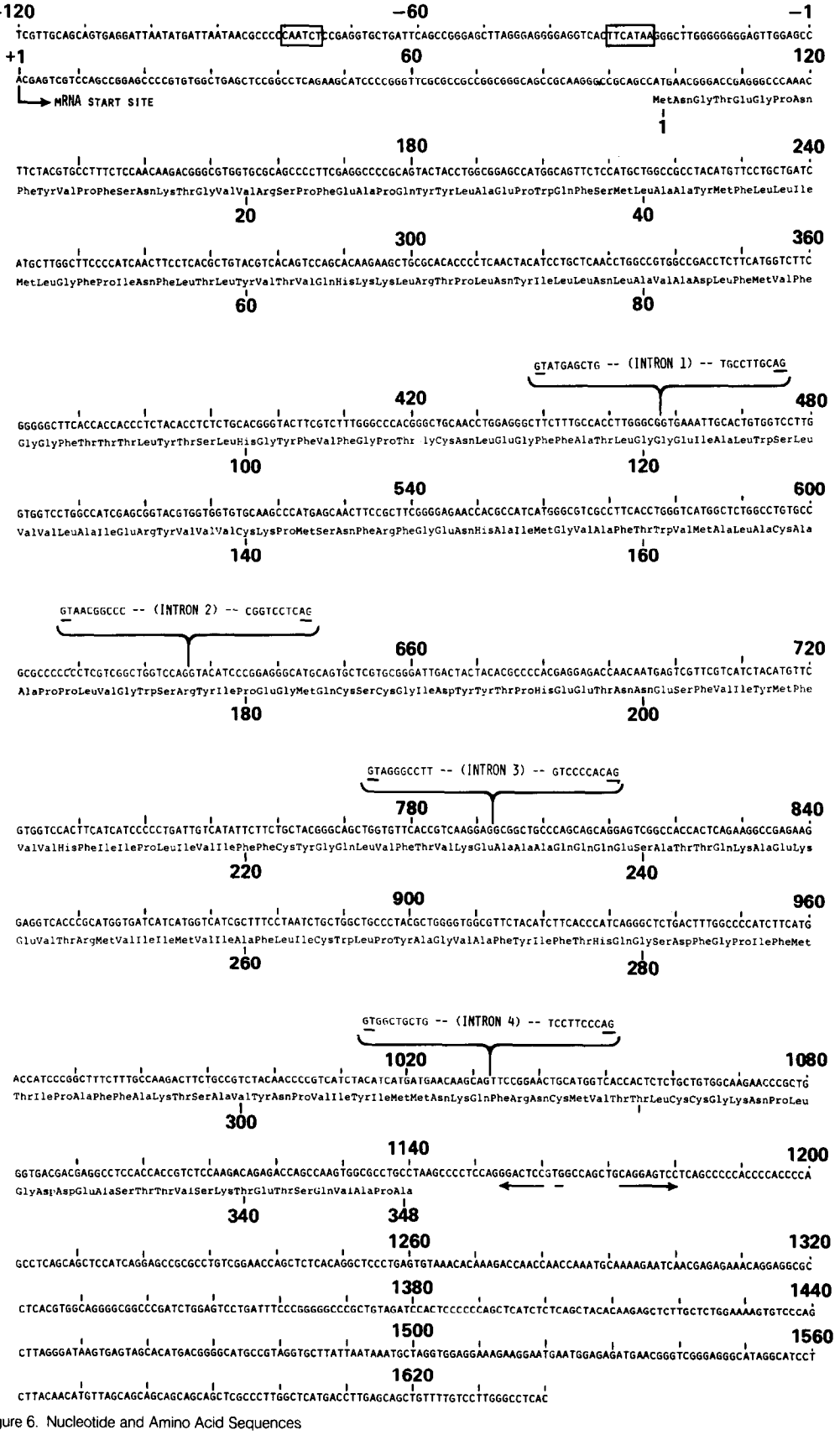

Figure 6. Nucleotide and Amino Acid Sequences

The sequence in the −120 to +1 region was determined from both strands of the Bd7 segment; one was 3′-end labeled (see Figure 5B), and the other 5′-end labeled at the unique Nco I site (see text, mRNA start site determination). Nucleotides in the exons are numbered +1 to 1640 according to their positions in the mRNA and were determined from both strands of the cDNA segments bd8 and bd20 and from one strand of Bd7, whereas the unnumbered intron sequences derive from only the one-strand Bd7 determination (Figures 5B, 5C). Exon lengths (nucleotides): exon 1 (5′-terminal), 457; exon 2, 169, exon 3, 166; exon 4, 240; exon 5, ~1520 (see text). See the text for the boxed sequences and the pair of inverted arrows.

prior to S1 nuclease digestion or cDNA extension by reverse transcriptase (see Muskavitch and Hogness, 1982, for a description of both methods). Gel electrophoresis of the extended and protected fragments, adjacent to a DNA sequence ladder of fragments that were also 5′-end labeled at and extend upstream from the Nco I site, revealed that their 3′ ends correspond, respectively, to positions +1 and −1 in Figure 6 (data not shown). We assume that the A at +1, rather than the C at −1, is the mRNA start site both because mRNA transcription is usually initiated with a purine and because the presumed 5′ cap structure in opsin mRNA might well extend the protection from S1 cleavage by one residue.

Analysis of the 5′-terminal 1639 bp in the 2.4 kb bd20 cDNA segment yielded the sequence occupying positions +2 to +1640 in Figure 6. This sequence represents all but the first of the 96 nucleotides in the 5′ untranslated region of the opsin mRNA, all of the 348 codons in its coding region, and 500 nucleotides of the 3′ untranslated region. Determination of the entire 1.5 kb bd8 cDNA sequence showed that it agreed exactly with the bd20 sequence in the region of their overlap (+154 to +1640).

The amino acid sequence generated by conceptual translation of the 348 codons (Figure 6) agrees with the partial opsin amino acid sequence of Hargrave et al. (1983) and is identical with the revised complete sequence reported by Ovchinnikov etal. (1982b), whose original sequence (Ovchinnikov et al., 1982a) lacked one member of the adjacent Tyr^{191}·Tyr^{192} pair. Given that the 348 codons begin with the first AUG triplet in the mRNA and are terminated by a UAA stop codon (Figure 6), it is evident that the opsin amino acid sequence derives from the primary translation product without loss or gain of any amino acids—a conclusion previously suggested by sequence analysis of the amino-terminal residues of an in vitro translation product (Schechter et al., 1979).

Sequence analysis of the mRNA homologous regions and adjacent DNA in the Bd7 genomic segment demonstrated that the coding region of the bovine opsin gene is interrupted by four introns (Figures 5B and 6), confirming electron microscopic observation of these introns in DNA heteroduplexes formed between the genomic and cDNA clones (data not shown). The cDNA sequence is exactly reproduced in the genomic DNA except that it is divided into the five exons defined by these introns. Figure 6 shows the sequences of the donor and acceptor splice junctions at the exon–intron boundaries. They are in good agreement with the consensus sequences for such junctions given by Breathnach and Chambon (1981; donor $5'\text{-}\frac{A}{C}AG\underline{GT}\frac{A}{G}AGT\text{-}3'$; acceptor $5'\text{-}\frac{CCCCCC}{TTTTTT}X\underline{CAG}\frac{G}{T}$, where the underlining indicates the invariant GT and AG dinucleotides). Because repetition of three to six nucleotides at the donor and acceptor sites for each intron prevents precise definition of the intron–exon boundaries, they were assigned according to the GT–AG rule, whereby the invariant GT of the donor and AG of the acceptor define the 5′ and 3′-terminal dinucleotides of each intron (Breathnach and Chambon, 1981).

The fifth exon appears to be the last in the opsin gene. The 3′ end of this exon was mapped by S1 nuclease protection of a genomic fragment strand that was 3′-end labeled at the Eco RI site within the fifth exon and extends 4.5 kb downstream to the Hind III site at the right end of the Bd7 segment (Figure 5B; Muskavitch and Hogness, 1982). Hybridization of this DNA to the opsin mRNA and subsequent S1 digestion yielded a 0.90 kb protected fragment. Given that that labeled Eco RI site is ~10 bp downstream of position 1640 (Figure 6), then the 3′ end of the fifth exon in the opsin mRNA is 2.55 kb from the 5′ end of that RNA. This distance is only ~0.10 kb less than the 2.65 kb observed for the length of the opsin mRNA (Figure 2 legend), leaving enough room for little else than the poly(A) tail. Given the identity of the 3′ ends of the opsin gene and its fifth exon, the gene is 6.4 kb long and contains a surprisingly long 1.4 kb 3′ untranslated region. Curiously, the 5′ end of this untranslated region is marked not only by a stop codon but also by an almost perfect 9 bp inverted repeat indicated by the two arrows in Figure 6.

The genomic sequence upstream from the mRNA start site (Figure 6) should contain a TATA or Goldberg–Hogness (M. L. Goldberg, Ph.D. thesis, Stanford University, 1979) box beginning at position −30 ± 4 and surrounded by GC-rich sequences (Breathnach and Chambon, 1981; Goldberg, op. cit). The best fit to these positional criteria and to the $TATA\frac{A}{T}A\frac{A}{T}$ consensus sequence (Breathnach and Chambon, 1981) is the boxed TTCATAA sequence shown in Figure 6. The other boxed sequence in Figure 6 indicates another region of sequence homology (the CAAT box) typically observed 70 to 80 bp upstream of the mRNA start site of eukaryotic genes (Breathnach and Chambon, 1981).

Discussion

By nucleotide sequence analysis of the cloned DNA segments, we have proved that they derive from the gene encoding bovine rhodopsin, deduced the complete amino acid sequence of this protein, mapped the intron–exon boundaries within this gene, and obtained the nucleotide sequence of part or all of its promoter region. We consider here the intron positions with reference to the general concept that introns mark the boundaries of structural or functional domains in the encoded protein and facilitate the evolutionary shuffling of such domains (Gilbert, 1978).

As an adjunct to a consideration of the secondary structure of bovine rhodopsin, we first compare it to the better defined structure of bacteriorhodopsin (reviewed by Henderson, 1977; Ovchinnikov, 1982; Engelman et al., 1982), which is also a membrane-embedded retinal chromoprotein. Bacteriorhodopsin acts as a light-driven proton pump in which photon absorption by, and consequent

isomerization of, its retinal chromophore induces protein conformational changes, as is the case for rhodopsin. In both proteins, the retinal is bound by a Schiff base linkage to a lysine located within the membrane just proximal to the hydrophilic C-terminal domain which resides on the cytoplasmic membrane surface (reviewed by Hargrave, 1982; Ovchinnikov, 1982). Because the N-terminal domain of both proteins is on the opposite membrane surface, both chains must traverse the membrane an odd number of times. That number is seven for bacteriorhodopsin, as shown by low resolution electron diffraction studies that indicate the molecule contains seven α-helical rods embedded in the membrane (Henderson and Unwin, 1975). A model specifying the amino acid sequences of these seven segments has been proposed by Engelman, Goldman, and Steitz (1982). In this model, the positions of the transmembrane segments are determined by the minima obtained when the calculated free energy cost of moving successive segments of the polypeptide chain from an aqueous to a nonpolar environment is plotted against the positions of these segments in the chain.

Figure 7 shows such a plot for bovine rhodopsin, where we have used a segment length of 21 residues, the minimum required for an α-helix to span the 30 Å thickness of the hydrophobic space in a lipid bilayer. The shape of this distribution is much like that obtained for bacteriorhodopsin using either 19- or 21-residue segments (Engelman et al., 1982; Steitz et al., 1982). When combined with the topological constraints imposed by the distribution of protease cleavage sites shown in Figure 7, it indicates that rhodopsin also contains seven transmembrane segments whose positions are given by the horizontal bars. Similar topographies for the rhodopsin transmembrane segments have been independently derived by Hargrave et al. (1983) and Ovchinnikov's group (Ovchinnikov, 1982).

Figure 7 also shows the positions at which the four introns interrupt the codon sequence. While the first intron does not appear to delineate the boundary of any domain that is recognizable in these preliminary structural models, the other three clearly do. Introns 2, 3, and 4 are alike in that they appear to mark the boundary between the carboxyterminus of a transmembrane segment and the amino terminus of an adjacent hydrophilic region, which either forms a membrane surface loop connecting two transmembrane segments or forms the C-terminal domain. This C-terminal domain is the site of light-dependent phosphorylation by rhodopsin kinase, the targets being the seven serine and threonine residues in the 334-to-343 region of the chain (reviewed by Hargrave, 1982). Recent studies in which the carboxy-terminal 12 residues were removed by selective proteolysis show that the activity of metarhodopsin II (one of the light-induced conformational isomers of rhodopsin) becomes insensitive to inhibition by rhodopsin kinase (Miller and Dratz, 1983). If phosphorylation is the mechanism of the observed ATP-dependent inactivation of metarhodopsin II activity, then the fifth exon may encode an activatable regulatory domain.

Examination of the positions assigned to the seven transmembrane segments in the above models reveals that the carboxy terminus of each hydrophobic segment is closely followed by a positively charged residue capable of interacting with the negatively charged phospholipid head groups. (Segment 1, $His^{65}Lys^{66}Lys^{67}$; 2, His^{100}; 3, Lys^{141}; 4, Arg^{177}; 5, Lys^{231}; 6, His^{278}; 7, Lys^{311}. Although the His^{100} lies just within segment 2 in Figure 7, it immediately follows that segment in the other two models; Ovchinnikov, 1982; Hargrave et al., 1983.) The positions of introns 2, 3, and 4 are such that the Arg^{177}, Lys^{231}, and Lys^{311} codons are the last or next-to-last codons in the exons encoding the 4th, 5th, and 7th transmembrane segments, respectively. This suggests the possibility that the carboxy-terminal border of the evolutionary elements of integral membrane proteins consists of a membrane-embedded hydrophobic segment followed by a positively charged residue.

The results reported here complete the first step in the

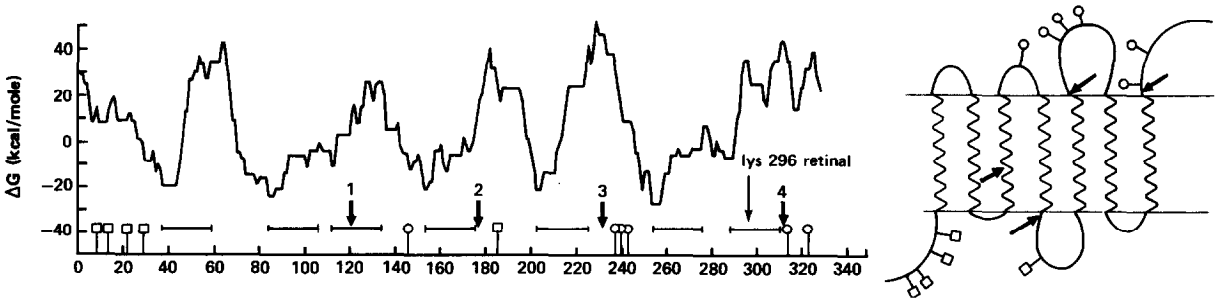

Figure 7. Topographic Correlations between Introns and Postulated Transmembrane Segments

The free energy of transfer for successive 21-amino-acid segments from water to nonpolar medium was calculated according to Steitz et al. (1982) and plotted against the position in the bovine rhodopsin chain of the first residue in each segment. The open circles and squares represent protease cleavage sites on the cytoplasmic and intradiscal membrane surfaces, respectively (Hargrave, 1982; Ovchinnikov, 1982; Mullen and Akhtar, 1983). The positions of the introns are indicated by the thick arrows labeled 1 to 4. The seven horizontal bars represent the postulated 21-residue transmembrane segments that occupy the hydrophobic space of the lipid bilayer, as indicated by the wavy lines in the schematic. Their positions were determined according to the following rules: All transmembrane segments must lie in the four intervals defined by the protease cleavage sites (interval I, 31–146; II, 147–186, III, 187–236; IV, 245–313). The number of segments per interval must be even if the sites defining an interval are on the same side of the membrane, and odd if on opposite sides. The position of the first residue in a segment must correspond to a minimum in the negative region of the ΔG plot, and the major minima in each interval must specify a segment.

plan outlined in the Introduction for the isolation of the human opsin genes. Experiments in which the coding sequence from the bovine opsin gene was used as a hybridization probe for the isolation of a human opsin gene indicate that the second step has also been successful.

Experimental Procedures

RNA Isolation

Fresh bovine eyes were obtained from Ferrara's Meat Company. Retinas were dissected within 2 hr and frozen on dry ice. Polysomes were isolated as described by Schechter (1974). Polysomal RNA was prepared by suspending the polysomes in 4.0 M guanidinium thiocyanate and sedimenting the RNA through a shelf of 5.7 M CsCl to form an RNA pellet, as described previously (Chirgwin et al., 1979). Poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography.

Construction of cDNA Libraries

Double-stranded S1-treated cDNA was synthesized as described by Buell et al. (1978) and Wickens et al. (1978). The cDNA was methylated with Eco RI methylase, the ends repaired with DNA polymerase I, and the cDNA ligated to synthetic Eco RI linkers (Maniatis et al., 1978). The linkered cDNA was cleaved with Eco RI and fractionated according to length on a 2 mm × 50 mm A-50M column. cDNAs longer than 500 bp were inserted into the unique Eco RI site in the *cI* gene of the λgt10 phage vector (T. Huynh and R. Davis, unpublished), a derivative of λ607 (Murray et al., 1977) with a smaller *b* region deletion. The DNA was packaged in vitro (Hohn and Murray, 1977) and plated onto Y1073, an *Hfl* derivative of C600 (Young and Davis, 1983). Phages bearing an insert at the Eco RI site are *cI*$^-$ and thus cannot lysogenize the *Hfl* host; phages lacking an insert are *cI*+ and lysogenize the *Hfl* host at such high frequency that they do not form plaques. Two micrograms of poly(A)$^+$ RNA yielded 2.5 × 10^5 independent clones of λ*cI*$^-$ recombinants.

Solid-Phase DNA Synthesis

The 15-mer DNA primer was synthesized by the solid-phase phosphotriester method (Miyoshi et al., 1980a, 1980b), with dimers, reagents, and a manual synthesizer purchased from Bachem, Inc. After removal of the protecting groups and cleavage from the resin, the 15-mer was purified by gel electrophoresis and its nucleotide sequence determined (Maxam and Gilbert, 1980).

Oligonucleotide-Primed cDNA Synthesis

For DNA sequence determination, the cDNA extension product was synthesized by mixing 0.05 μg of the 15-mer DNA, 5′- end ^{32}P-labeled to 0.7 × 10^9 cpm/μg (Maxam and Gilbert, 1980), with 30 μg of retinal polysomal poly(A)$^+$ RNA in a total volume of 10 μl of 10 mM Tris-HCl (pH 7.5), 1.0 mM EDTA. After heating this solution for 1.0 min at 75°C, 1.0 μl of 3 M KCl was added, and the sample incubated for 15 min at 30°C. cDNA synthesis was initiated by the addition of 50 μl of 60 mM Tris-HCl (pH 8.3), 40 mM KCl, 10 mM $MgCl_2$, 0.5 mM DTT, 1 mM of each deoxynucleotide triphosphate, and 20 units of AMV reverse transcriptase. After a 10 min incubation at 30°C, the sample was placed at 37°C for 1.0 hr. Gel electrophoresis of the product under denaturing conditions revealed a number of discrete ^{32}P-labeled cDNAs up to but not exceeding 1 kb in length. The product was heat-denatured and fractionated on an A-50M sizing column, and fractions containing cDNAs ≥ 0.2 kb (1.5 × 10^5 cpm) were pooled. The radioactive cDNA extension product used for plaque screening was prepared under identical conditions, except that α-^{32}P-dCTP (800 Ci/mmole; 2 mCi/ml) was used in place of unlabeled dCTP.

Oligonucleotide Hybridization to RNA Blots

RNA was treated with glyoxal (McMaster and Carmichael, 1977), electrophoresed, transferred to diazatized paper (Alwine et al., 1977), and hybridized with the 15-mer DNA primer that had been ^{32}P-labeled (0.7 × 10^9 cpm/μg) at its 5′ end by polynucleotide kinase (Maxam and Gilbert, 1980). The hybridization was carried out at 23°C for 65 hr in 1.0 M NaCl, 50 mM sodium phosphate (pH 7.0), 5.0 mM EDTA, and 5 × 10^6 cpm/ml of labeled oligonucleotide. The paper was washed at 8°C for 30 min with five changes of 1.0 M NaCl, 50 mM sodium phosphate (pH 7.0), 5.0 mM EDTA, and 0.1% (w/v) Sarkosyl and exposed to Kodak XAR-5 film at −80°C in the presence of a Du Pont Lightning Plus intensifying screen.

Hybridization to DNA

Hybridizations were carried out at 42°C in 50% formamide, 1.0 M NaCl, 50 mM sodium phosphate (pH 7.0), and 5 mM EDTA and the filters were washed at 42°C with 20 mM NaCl, 1.0 mM sodium phosphate (pH 7.0), 5 mM EDTA.

Construction of Genomic DNA Libraries Enriched for Opsin Genes

Bovine genomic DNA was isolated from retinal nuclei by equilibrium sedimentation in CsCl gradients (Davis et al., 1980). Hind III or Eco RI digests of 100 μg of this DNA were fractionated acording to fragment length by electrophoresis in 0.4% SeaPlaque agarose gels (F.M.C. Corp., Marine Colloids Division). The gel was then equilibrated with 0.3 M NaCl and cut into slices, and the DNA from individual slices was isolated by melting the gel at 70°C for 15 min., extracting twice with warm phenol, extracting once with ether, and precipitating the DNA by the addition of two volumes of ethanol. Aliquots of DNA from the slices were electrophoresed, transferred to diazatized paper (Alwine et al., 1977), and hybridized with ^{32}P-labeled (Rigby et al., 1977) cloned opsin cDNA to determine which samples contained sequences from the opsin gene. The DNA in the selected samples was cloned in λ762 (for Hind III inserts; Murray et al., 1977) or Charon 4 (for Eco RI inserts; Blattner et al., 1977) phage vectors. This electrophoretic fractionation provided an enrichment of 20-fold over total genomic DNA in clones containing opsin genes.

DNA Sequence Analysis

DNA nucleotide sequences were determined by the method of Maxam and Gilbert (1980), using 40 and 80 cm sequencing gels.

Acknowledgments

We thank Roberto Kolter and Charles Yanofsky for their generous help in preparing the synthetic oligodeoxynucleotide, Bernard Fung and Lubert Stryer for their generous gift of rabbit anti-rhodopsin sera, and Wolfgang Baehr, Meridithe Applebury, Arnold Satterthwait, Daniel Oprian, and H. Gobind Khorana for their generous communication of results prior to publication. We also thank Lubert Stryer for his advice and encouragement. J. N. is supported by a fellowship from the Medical Scientist Training Program of the National Institutes of Health, and this research was supported by a National Institutes of Health grant to D. S. H.

Received August 2, 1983

References

Alwine, J. C., Kemp, D. J., and Stark, G. R. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyoxymethyl-paper and hybridization with DNA probes. Proc. Nat. Acad. Sci. USA *74*, 5350–5354.

Baehr, W., and Applebury, M. (1983). Characterization of bovine rhodopsin mRNA and cDNA. Biophys. J. *41*, 340a.

Benton, W. D., and David, R. W. (1977). Screening lambda gt recombinant clones by hybridization to single plaques *in situ*. Science *196*, 180–182.

Blattner, F. R., Williams, B. G., Blechl, A. E., Thompson, K. D., Faber, H. E., Furlong, L. A., Grunwald, D. J., Kefier, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L., and Smithies, O. (1977). Charon phages: safer derivatives of bacteria-phage lambda for DNA cloning. Science *196*, 161–169.

Boynton, R. (1979). Human Color Vision. (New York: Holt, Reinhart and Winston).

Breathnach, R., and Chambon, P. (1981). Organization and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem. *50*, 349–383.

Buell, G. N., Wickens, M. P. Payvar, F., and Schimke, R. T. (1978). Synthesis of full length cDNAs from four partially purified oviduct mRNAs. J. Biol. Chem. *253*, 2471–2482.

Chirgwin, J. M. Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18*, 5294–5299.

Davis, R. W., Botstein, D., and Roth J. R. (1980). Advanced Bacterial Genetics. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 116–119.

Engelman, D. M., Goldman, A., and Steitz, T. A. (1982). The identification of helical segments in the polypeptide chain of bacteriorhodopsin. Meth. Enzymol. *88*, 81–88.

Gilbert, W. (1978). Why genes in pieces? Nature *271*, 501.

Gough, N. M., and Adams, J. R. (1978). Immunoprecipitation of specific polysomes using Staphylococcus aureus: purification of the immunoglobulin κ chain messenger RNA from the mouse myeloma MPC11. Biochemistry *17*, 5560–5566.

Hargrave, P. A. (1982). Rhodopsin chemistry, structure and topography. In Progress in Retinal Research, N. Osborne and G. Chader, eds. (Oxford: Pergamon), pp. 1–51.

Hargrave, P. A., McDowell, J. H., Siemiatkowski-Juszczak, E. C., Fong, S.-L., Kuhn, H., Wang, J. K., Curtis, D. R., Mohana Rao, J. K., Argos, P., and Feldmann, R. J. (1982). The carboxy terminal one-third of bovine rhodopsin: its structure and function. Vision Res. *22*, 1429–1438.

Hargrave, P. A., McDowell, J. H., Curtis, D. R., Wang, J. K., Juszczak, E., Fong, S.-L., Mohana Rao, J. K., and Argos, P. (1983). The structure of bovine rhodopsin. Biophys. Struct. Mech. *9*, 235–244.

Henderson, R. (1977). The purple membrane from Halobacterium halobium. Ann. Rev. Biophys. Bioeng. *6*, 87–109.

Henderson, R., and Unwin, P. N. T. (1975). Three dimensional model of purple membrane obtained by electron microscopy. Nature *257*, 28–32.

Hohn, B., and Murray, K. (1977). Packaging recombinant DNA molecules into bacteriophage particles in vitro. Proc. Nat. Acad. Sci. USA *74*, 3259–3263.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., and Efstratiadis, A. (1978). The isolation of structural genes from libraries of eucaryotic DNA. Cell *15*, 687–701.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

McMaster, G. K., and Carmichael, G. G. (1977). Analysis of single and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Nat. Acad. Sci. USA *74*, 4835–4838.

Miller, J. L., and Dratz, E. A. (1983). Phosphorylation at sites near rhodopsin's carboxy-terminus regulates light initiated cyclic-GMP hydrolysis. Biophys. J. *41*, 343a.

Miyoshi, K., Miyaka, T., Hozumi, T., and Itakura, K. (1980a). Solid phase synthesis of polynucleotides. II. Synthesis of polythymidylic acid by the block coupling phosphotriester method. Nucl. Acids Res. *8*, 5473–5490.

Miyoshi, K., Huang, T., and Itakura, K. (1980b). Solid phase synthesis of polynucleotides. III. Synthesis of polynucleotides with defined sequences by the block coupling phosphotriester method. Nucl. Acids Res. *8*, 5491–5505.

Mullen, E., and Akhtar, M. (1983). Structural studies on membrane-bound bovine rhodopsin. Biochem. J. *211*, 45–54.

Murray, N. E., Brammer, W. J., and Murray, K. (1977). Lambdoid phages that simplify the recovery of in vitro recombinants. Mol. Gen. Genet. *150*, 53–61.

Muskavitch, M. A. T., and Hogness, D. S. (1982). An expandable gene that encodes a Drosophila glue protein is not expressed in variants lacking remote upstream sequences. Cell *29*, 1041–1051.

Ovchinnikov, Y. A. (1982). Rhodopsin and bacteriorhodopsin: structure–function relationships. FEBS Lett. *148*, 179–191.

Ovchinnikov, Y. A., Abdulaev, N. G., Feigina, M. Y., Artamonov, I. D., Zolotarev, A. S., Miroshnikov, A. I., Martynov, V. I., Kustina, M. B., Kudelin, A. B., and Bogachuk, A. S. (1982a). The complete amino acid sequence of visual rhodopsin. Bioorg. Khim. *8*, 1011–1014.

Ovchinnikov, Y. A., Abdulaev, N. G., Feigina, M. Y., Artamonov, I. D., Zolotarev, A. S., Miroshnikov, A. I., Martynov, V. I., Kostina, M. B., Kudelin, A. B., and Bogachuk, A. S. (1982b). Bioorg. Khim. *8*, 1424–1427.

Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. *113*, 237–251.

Schechter, I. (1974). Use of antibiodies for the isolation of biologically pure messenger ribonucleic acid from fully functional eukaryotic cells. Biochemistry *13*, 1857–1885.

Schechter, I., Burstein, Y., Zemell, R., Ziv, E., Kantor, F., and Papermaster, D. (1979). Messenger RNA of opsin from bovine retina: isolation and partial sequence of the in vitro translation product. Proc. Nat. Acad. Sci. USA *76*, 2654–2658.

Steitz, T. A., Goldman, A., and Engelman, D. M. (1982). Quantitative application of the helical hairpin hypothesis to membrane proteins. Biophys. J. *37*, 124–125.

Sutcliffe, J. G. (1978). pBR322 restriction map marked from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. *5*, 2721–2728.

Wald, G. (1968). The molecular basis of visual excitation. Nature *219*, 800–807.

Wickens, M. P., Buell, G. N., and Schimke, R. T. (1978). Synthesis of double-stranded DNA complementary to lysozyme, ovomucoid, and ovalbumin mRNAs. J. Biol. Chem. *253*, 2483–2495.

Young, R. A., and Davis, R. W. (1983). Efficient isolation of genes using antibody probes. Proc. Nat. Acad. Sci. USA *80*, 1194–1198.