The Journal of Neuroscience, May 1987, 7(5): 1558–1566

A Second Opsin Gene Expressed in the Ultraviolet-Sensitive R7 Photoreceptor Cells of *Drosophila melanogaster*

Craig Montell, Kevin Jones, Charles Zuker[a], and Gerald Rubin

Department of Biochemistry, University of California, Berkeley, California 94720

We have identified a second *Drosophila* opsin gene, Rh4, which is expressed specifically in the ultraviolet-sensitive R7 photoreceptor cells. The two R7 opsins, Rh3 and Rh4, are expressed in nonoverlapping subsets of R7 cells. The Rh4 opsin shares 72% amino acid identity with Rh3, but is only about 35% homologous with the other 2 *Drosophila* opsins, *ninaE* and Rh2. Twenty-eight amino acids are conserved among all 4 invertebrate and 5 vertebrate opsin genes thus far sequenced, suggesting that these residues play an important role in rhodopsin function.

The compound eye of the fruit fly *Drosophila melanogaster* contains approximately 800 repeat units, or ommatidia. Each ommatidium contains 8 photoreceptor cells, which fall into 3 classes based on their topological arrangement and spectral properties (reviewed by Pak and Grabowski, 1978; Hardie, 1983). The major class consists of the 6 outer photoreceptor cells, R1–6, which extend the length of the retina and express a rhodopsin that absorbs maximally at 480 nm (Ostroy et al., 1974). The 2 minor classes of photoreceptor cells, R7 and R8, occupy the distal and proximal central regions of the ommatidia, respectively. The R8 cells respond maximally to blue light and the R7 photoreceptor cells to UV light (Harris et al., 1976).

The major visual pigment is encoded by the genetically identified *ninaE* locus (O'Tousa et al., 1985; Zuker et al., 1985). The photopigments found in the R7 and R8 photoreceptor cells are not affected by mutations in the *ninaE* locus and thus must be encoded by other genes. Indeed, an opsin gene, Rh2, expressed specifically in photoreceptor cell R8 has been identified (Cowman et al., 1986), and in the accompanying paper (Zuker et al., 1987) we report the isolation of an opsin, Rh3, expressed in photoreceptor cell R7. Microspectrophotometric studies carried out mainly in the larger flies, *Musca* and *Calliphora* (reviewed in Hardie, 1983), have suggested the existence of 2 distinct R7 photopigments. To determine whether a second opsin gene is expressed in photoreceptor cell R7 of *Drosophila,* we used an Rh3 DNA clone to screen a *Drosophila* DNA library for related opsin genes. We report here the isolation and characterization of a second opsin gene, Rh4, expressed specifically in photoreceptor cell R7. We also demonstrate that expression of the Rh3 and Rh4 genes is spatially restricted to nonoverlapping subsets of R7 cells. Together, these 4 *Drosophila* rhodopsin genes may account for all the spectral properties of the *Drosophila* compound eye.

Received Sept. 26, 1986; accepted Nov. 25, 1986.

This work was supported by grants from the NIH to G. R. C. M. is supported by a NIH postdoctoral fellowship, K.J. by a NSF predoctoral fellowship, and C.Z. by a Jane Coffin Childs postdoctoral fellowship. We thank Drs. Daniel Kalderon, Kevin Moses, and Jeremy Nathans for improving this manuscript with their helpful comments.

Correspondence should be addressed to Craig Montell at the above address.

[a] Present address: Department of Biology, B-022, University of California at San Diego, La Jolla, CA 92093.

Copyright © 1987 Society for Neuroscience 0270-6474/87/051558-09$02.00/0

Materials and Methods

Isolation of Rh4 cDNA and genomic λ-phage clones. A *Drosophila* λgt10 complementary DNA (cDNA) library, prepared from Oregon R adult head poly(A)$^+$ RNA (a gift from B. Yedvobnick and S. Artavanis-Tsakonas), was screened with p512SB1 nick-translated with α-^{32}P-dNTPs. The plasmid p512SB1 was constructed by subcloning a 2.9 kbase *Sal* I-*Bam* HI Rh3 genomic fragment from λ512 (Levy et al., 1982) to pSP64 (Promega Biotech, Madison, WI). This 2.9 kbase *Sal* I-*Bam* HI fragment corresponds to the *Sal* I-*Bam* HI fragment in λDmRh3 (Zuker et al., 1987). The filters were hybridized at 65°C in 5× SSCP (750 mM NaCl, 100 mM NaH_2PO_4, pH 6.8, 75 mM sodium citrate), 5× Denhardt's solution (1× Denhardt's solution is 0.02% PVP-40, 0.02% bovine serum albumin, 0.02% Ficoll), 0.5% SDS, and 100 μg/ml salmon sperm DNA, and washed at 65°C in 2× SSC (1× SSC is 150 mM NaCl, 15 mM sodium citrate). One of the positive clones, λc512-2, contained a 1.4 kbase insert DNA fragment that mapped via *in situ* hybridization to polytene salivary gland chromosomes, to the cytogenetic position 73D on the third chromosome.

An EMBL4 (Frischauf et al., 1983) *Drosophila* Canton S genomic DNA library (the gift of V. Pirrotta) was screened with a *Pst* I-*Bam* HI c512-2 DNA fragment (nucleotides 1100-1402), nick-translated with α-^{32}P-dNTPs. The filters were hybridized at 65°C in buffer containing 2× SSCP, 4× Denhardt's, 0.5% SDS, and 100 μg/ml salmon sperm DNA, and were washed at 65°C in 0.1× SSC. Three positive clones were isolated and each hybridized to polytene band 73D.

In situ *hybridization to polytene chromosomes. In situ* hybridization to polytene chromosomes was carried out as described by Zuker et al. (1985).

DNA sequencing. DNA sequencing was carried out according to the dideoxy chain termination method (Sanger et al., 1977) using ^{35}S-dATP as the radioactive nucleotide (Bankier and Barrell, 1983). The plasmid pc512-2 was constructed by subcloning the 1.4 kbase *Eco* RI cDNA

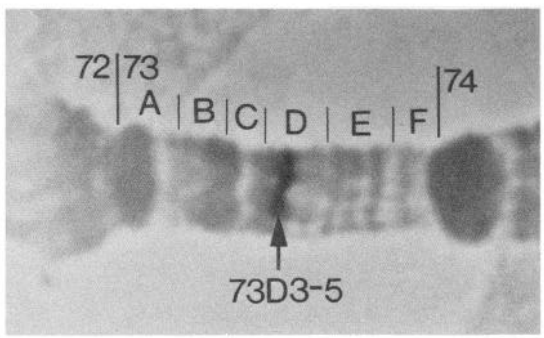


Figure 1. In situ hybridization to Canton S salivary gland polytene chromosomes. The plasmid, pc512-2, was biotinylated by nick-translation with Bio-16-dUTP and used as a hybridization probe to determine the chromosomal location of Rh4. Shown is the 73 region of chromosome 3. The *arrow* indicates the site of hybridization at 73D3-5.

```
                                                          ↓                    *
GTTGGCAGCA CAAAATGCGA TATAAACGAG GCATTGCGGT GATTGGAAAC     AGTTCAGCCC TGTGGCGCAC ACAG        24

AGCGAAACGG GTAGCGGTTA AGGCGTATCG GTTGGGAATC GGTTCGGTAT CGGGTTGACC GAT ATG GAG CCG TTG       99
                                                                    MET Glu Pro Leu         4

TGC AAC GCC AGT GAA CCG CCA CTT CGC CCG GAA GCG CGG AGC TCT GGC AAC GGA GAC TTA CAG TTC    165
Cys Asn Ala Ser Glu Pro Pro Leu Arg Pro Glu Ala Arg Ser Ser Gly Asn Gly Asp Leu Gln Phe     26

CTC GGA TGG AAT GTG CCG CCG GAT CAG ATT CAA TAC ATT CCG GAG CAC TGG CTG ACC CAG CTG GAA    231
Leu Gly Trp Asn Val Pro Pro Asp Gln Ile Gln Tyr Ile Pro Glu His Trp Leu Thr Gln Leu Glu     48

CCG CCC GCG TCC ATG CAC TAC ATG CTG GGC GTC TTC TAC ATA TTT CTC TTC TGC GCC TCG ACA GTG    299
Pro Pro Ala Ser MET His Tyr MET Leu Gly Val Phe Tyr Ile Phe Leu Phe Cys Ala Ser Thr Val     70

GGT AAT GGC ATG GTC ATC TGG ATC TTC AGC ACA TCC AAG TCG CTG AGG ACA CCA TCC AAT ATG TTC    363
Gly Asn Gly MET Val Ile Trp Ile Phe Ser Thr Ser Lys Ser Leu Arg Thr Pro Ser Asn MET Phe     92

GTG CTG AAT CTG GCC GTC TTC GAT CTG ATC ATG TGC CTC AAG GCG CCG ATC TTC AAC AGC TTC CAT    429
Val Leu Asn Leu Ala Val Phe Asp Leu Ile MET Cys Leu Lys Ala Pro Ile Phe Ile Tyr Asn Ser    114

CGA GGA TTC GCC ATC TAC CTG GGC AAT ACC TGG TGC CAG ATA TTC GCC TCC ATT GGC TCC TAT TCG    495
Phe His Arg Gly Phe Ala Leu Gly Asn Thr Trp Cys Gln Ile Phe Ala Ser Ile Gly Ser Tyr Ser    136

GGT ATT GGT GCT GGT ATG ACG AAC GCG GCC ATA GGA TAC GAT CGA TAC AAT GTG ATC ACC AAG CCC    561
Gly Ile Gly Ala Gly MET Thr Asn Ala Ala Ile Gly Tyr Asp Arg Tyr Asn Val Ile Thr Lys Pro    158

ATG AAC CGC AAC ATG ACC TTC ACC AAG GCG GTG ATA ATG AAT ATA ATC ATC TGG TTG TAC TGC ACA    627
MET Asn Arg Asn MET Thr Phe Thr Lys Ala Val Ile MET Asn Ile Ile Ile Trp Leu Tyr Cys Thr    180

                                  GTGAGATTGT-9.0 kb-GCATTTCCAG
                                  |

CCG TGG GTT GTC CTG CCG CTA ACC CAG TTC TGG GAT CGA TTC GTG CCA GAG GGC TAC CTC ACG TCC    693
Pro Trp Val Val Leu Pro Leu Thr Gln Phe Trp Asp Arg Phe Val Pro Glu Gly Tyr Leu Thr Ser    202

TGC TCC TTC GAC TAT CTT TCG GAC AAC TTT GAC ACC CGG TTG TTT GTG GGC ACC ATC TTC TTT TTC    759
Cys Ser Phe Asp Tyr Leu Ser Asp Asn Phe Asp Thr Arg Leu Phe Val Gly Thr Ile Phe Phe Phe    224

AGC TTC GTG TGT CCC ACG CTG ATG ATC CTT TAC TAC TAC TCG CAG ATC GTG GGC CAT GTC TTC AGC    825
Ser Phe Val Cys Pro Thr Leu MET Ile Leu Tyr Tyr Tyr Ser Gln Ile Val Gly His Val Phe Ser    246

CAC GAA AAG GCC CTA CGG GAG CAG GCC AAG AAA ATG AAC GTG GAG TCG CTG CGC TCC AAT GTG GAC    891
His Glu Lys Ala Leu Arg Glu Gln Ala Lys Lys MET Asn Val Glu Ser Leu Arg Ser Asn Val Asp    268

AAG AGC AAG GAG ACG GCG GAG ATA CGG ATT GCG AAG GCG GCT ATC ACC ATC TGC TTC CTG TTC TTC    957
Lys Ser Lys Glu Thr Ala Glu Ile Arg Ile Ala Lys Ala Ala Ile Thr Ile Cys Phe Leu Phe Phe    290

GTG TCG TGG ACG CCC TAC GGC GTA ATG TCG CTG ATC GGG GCA TTC GGG GAT AAG AGT CTG CTT ACA   1023
Val Ser Trp Thr Pro Tyr Gly Val MET Ser Leu Ile Gly Ala Phe Gly Asp Lys Ser Leu Leu Thr    312

CAA GGA GCC ACG ATG ATC CCG GCC TGC ACC TGC AAA CTG GTG GCG TGC ATA GAC CCA TTC GTC TAT   1089
Pro Gly Ala Thr MET Ile Pro Ala Cys Thr Cys Lys Leu Val Ala Cys Ile Asp Pro Phe Val Tyr    334

GCC ATA AGT CAC CCC AGA TAC CGC TTG GAG CTG CAG AAG CGC TGT CCC TGG CTG GGA GTC AAC GAA   1155
Ala Ile Ser His Pro Arg Tyr Arg Leu Glu Leu Gln Lys Arg Cys Pro Trp Leu Gly Val Asn Glu    356

AAG TCT GGG GAG ATC TCT TCG GCG CAG TCC ACG ACC ACC CAG GAG CAG CAA CAG ACT ACC GCT GCA   1221
Lys Ser Gly Glu Ile Ser Ser Ala Gln Ser Thr Thr Thr Gln Glu Gln Gln Gln Thr Thr Ala Ala    378

TAGAACCAAG GACAACTCTA CTCTAAGACA ACTGACCATG TAACATGAAA GCCAAGGAAA AAGTATAAAA TGCCGACAAC
GAAACTGTAT AACATTAATT TTATAATTTG TAGTGTGACA TTCTTGAGTT TGAAATAAAT AAATAGTAAC TTATTGCAAA
CGAAGTAGAA AATGAAAAAA AAAAAAAAAA AAAAAAAAAA AA                                          1396
```

Figure 2. DNA and amino acid sequence of Rh4. The *top line* lists in *italics* the genomic DNA sequence immediately 5′ to the beginning of the Rh4 RNA. The transcription initiation site is indicated by the *arrow* and the first nucleotide in the Rh4 cDNA, c512-2 (+22) is identified by the *asterisk.* The first nucleotide in the c512-2 poly(A) tract is *1396.* The cDNA and flanking genomic DNA were sequenced entirely on both strands. The deduced amino acid sequence is shown *below the DNA sequence.* The *column on the right* lists the running tally of nucleotides and amino acids encoded in Rh4. The genomic DNA sequence at the 5′ and 3′ ends of the 9.0 kbase intron is shown *above the nucleotides 676 and 677,* which are separated by this intron. The TATA box homology and the 10 nucleotide sequence encompassing the 3′ consensus sequence, AATAAA, are *underlined.*

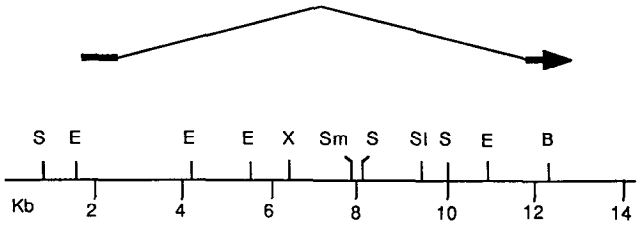


Figure 3. Physical map of Rh4. The *bottom line* represents the genomic DNA encompassing the Rh4 coding region demarcated in kilobase pairs (*Kb*). The location of the restriction sites *Bgl* II (*B*), *Eco* RI (*E*), *Sac* II (*S*), *Sal* I (*Sl*), *Sma* I (*Sm*), and *Xho* I (*X*) are indicated. The 2 Rh4 exons are depicted by *bold lines*. The direction of transcription is indicated by the *arrowhead*. To determine the regions of genomic DNA coding for Rh4, λRh4-2 was digested with various restriction endonucleases, transferred to nitrocellulose, and probed with the entire Rh4 cDNA, a 5′ end-specific probe and a 3′ end-specific probe. Two noncontiguous segments of genomic DNA homologous to the Rh4 cDNA probe were sequenced and the structure of the Rh4 gene was obtained by comparing the cDNA sequence with the genomic sequence.

fragment from λc512-2 into the *Eco* RI site of pEMBL9 (Dente et al., 1983). The Rh4 genomic plasmid clones pRh4-E1 and pRh4-E2 were constructed by subcloning the *Eco* RI genomic fragments encoding the 3′ (11.0–14.2 kbase) and 5′ (1.6–4.2 kbase) exons from λRh4-2 (see Fig. 4) into the *Eco* RI sites of pUC13 and pEMBL9, respectively. Ten micrograms of the plasmids pc512-2, pRh4-E1, and pRh4-E2 were randomly sheared by sonication and the ends were repaired with T4 DNA polymerase and Klenow for 4 hr in T4 polymerase buffer (Maniatis et al., 1982). Fragments in the 300–600 nucleotide size range were electroeluted from an agarose gel, inserted into the *Sma* I site of M13 mp10 by ligation for 48 hr at 4°C and introduced into the *Escherichia coli* strain TG1 (the gift of Toby Gibsom, Medical Research Council, Cambridge, England). Recombinant M13 clones containing Rh4 cDNA fragments or genomic inserts encoding the 5′ or 3′ exons were identified by plaque hybridization to the gel purified 1.4 kbase pc512-2 cDNA fragment nick-translated with α-^{32}P-dNTPs.

RNA blot analysis. Polyadenylated RNA was prepared as described (Montell et al., 1985) from the Oregon R strain of *D. melanogaster* and from *sev*^{LY3} flies and fractionated on 3% formaldehyde, 1.5% agarose gels as previously described (Maniatis et al., 1982). The RNAs were transferred to nitrocellulose and then hybridized to the *Pst* I-*Bam* HI c512-2 DNA fragment (nucleotides 1100–1402), nick-translated with α-^{32}P-dNTPs. The hybridization was at 65°C in 2× SSCP, 2× Denhardt's, 0.5% SDS, 250 μg/ml salmon sperm DNA, and the filter was washed at 65°C in 0.1× SSC.

In situ hybridization to tissue sections. Frozen 8 μm sections of adult fly heads were prepared and hybridizations of ^{3}H-labeled Rh3 [0.9 kbase *Pst* I-Hind III genomic fragment, nucleotides 1070–2016 (Zuker et al., 1987)] and Rh4 (*Pst* I-*Eco* RI cDNA fragment, nucleotides 1100–1402)-specific probes were as described (Hafen et al., 1983), with the omission of the acid and pronase treatments.

Results

Isolation of a cDNA related to Rh3

A cDNA library made from mRNAs expressed in the heads of adult *Drosophila* was screened for sequences homologous to a 2.9 kbase *Sal* I-*Bam* HI genomic fragment encompassing the Rh3 opsin gene. In addition to isolating some cDNAs that

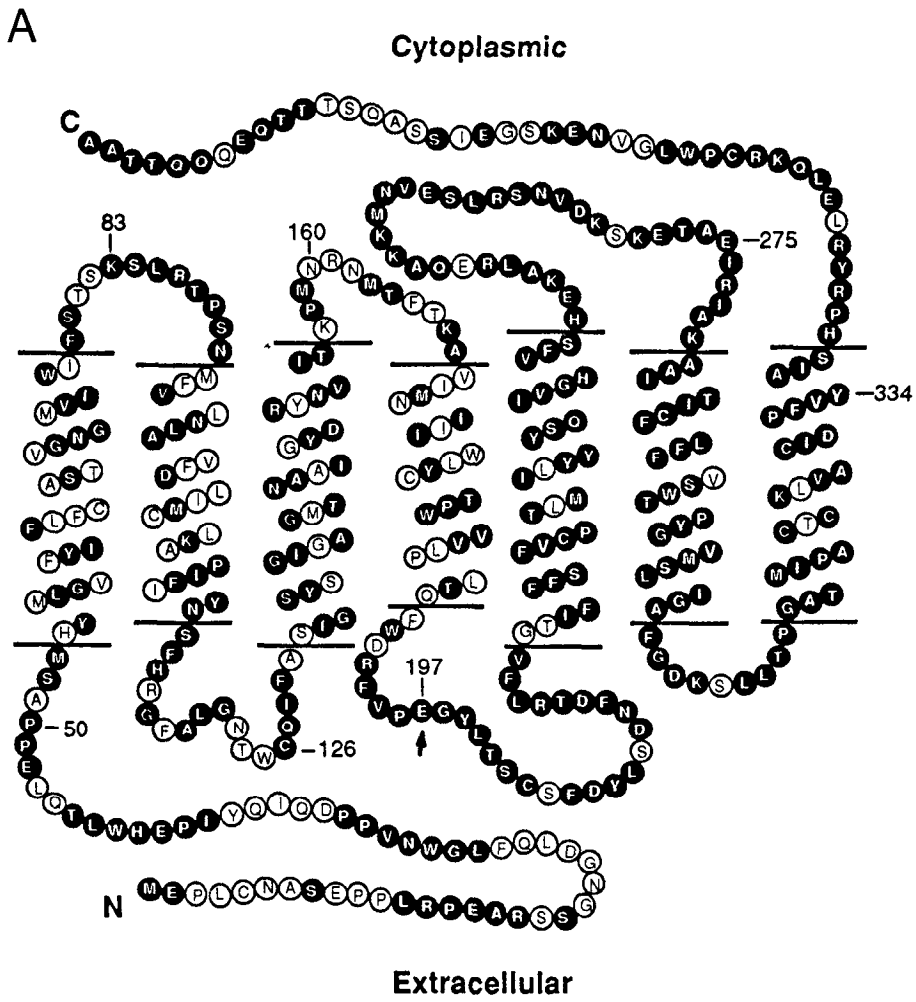


Figure 4. Locations of amino acid identities between Rh4 and other opsins. Shown is the proposed structure of the Rh4 opsin according to models previously presented for other opsins (Ovchinnikov, 1982; Hargrave et al., 1983; Nathans and Hogness, 1983, 1984; O'Tousa et al., 1985; Zuker et al., 1985, 1987; Cowman et al., 1986; Nathans et al., 1986). An adjustment was made relative to the previously proposed structures for *ninaE* and Rh2 so as to maintain the lengths of the second cytoplasmic and extracellular loops with the vertebrate rhodopsins. The 7 transmembrane segments correspond to regions having positive hydrophobicity indices, according to the algorithm of Kyte and Doolittle (1982). The *N*- and *C*-termini of Rh4 are indicated. Amino acid identities are indicated by *shaded circles*. The amino acids are represented by IUPAC-IUB CBN one letter abbreviations. The *arrow* in *A* indicates the glutamic acid residue (amino acid 197) interrupted by the 9.0 kbase intron. *A*, Comparison among the *Drosophila* R7 opsins, Rh3 and Rh4. *B*, Comparison among the 4 *Drosophila* rhodopsins, bovine rhodopsin, human rhodopsin, and the 3 human cone visual pigments. The methionine (*M*, residue 1) at the N-termini of the opsins is not shaded in *B*. The complete aligned amino acid sequences of these opsins are shown in figure 2 of the accompanying paper (Zuker et al., 1987).

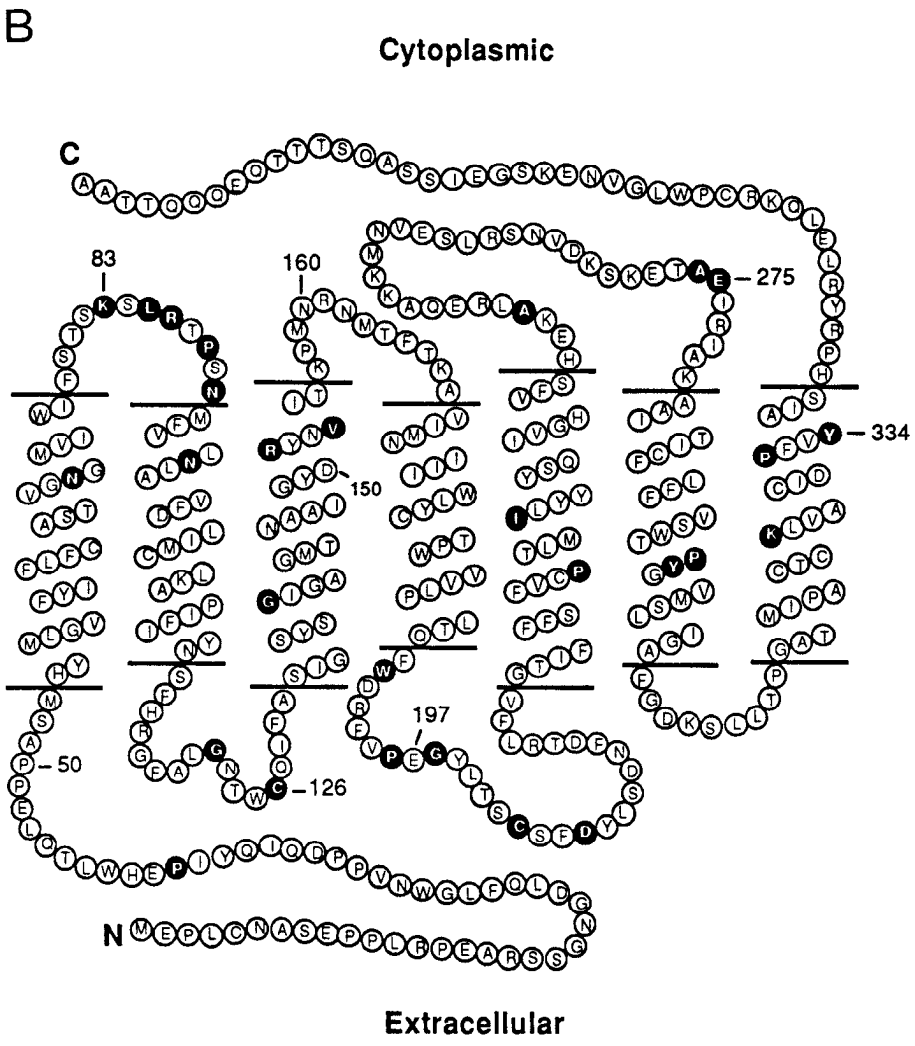
mapped, by *in situ* hybridization, to polytene salivary gland chromosome position 92D, the cytogenetic location of Rh3 (Zuker et al., 1987), we isolated a 1.4 kbase cDNA, λc512-2, which mapped to 73D3-5 on the third chromosome (Fig. 1). No mutation affecting vision that maps to this position has been reported.

cDNA clone mapping to 73D encodes an opsin

Figure 2 displays the complete DNA sequence of this cDNA, as well as that of the 378-residue polypeptide it is predicted to encode. This polypeptide is highly homologous to the other *Drosophila* opsins (see below), and shares several features common to rhodopsin molecules from a variety of organisms. These include 7 potential transmembrane domains, a putative retinal binding site (lysine residue 324) in the seventh transmembrane domain, a site for N-linked glycosylation near the N-terminus, and several serine and threonine residues near the C-terminus that may serve as sites for light-dependent phosphorylation (Ovchinnikov, 1982; Hargrave et al., 1983; Nathans and Hogness, 1983, 1984; O'Tousa et al., 1985; Zuker et al., 1985, 1987; Cowman et al., 1986; Nathans et al., 1986). On the basis of these criteria, as well as its pattern of expression (see below), we conclude that the DNA sequence corresponding to c512-2 encodes an opsin. This gene will be referred to as Rh4.

Rh4 is encoded in 2 exons

In order to determine the structure of the Rh4 gene, we isolated a genomic clone encoding the Rh4 opsin. A bacteriophage λ-library was screened for sequences homologous to a 302 bp DNA fragment encoding the last 33 amino acids and the 3′ untranslated region of Rh4. This sequence does not cross-hybridize to the other 3 *Drosophila* opsin genes under the hybridization conditions used. A physical map of the Rh4 genomic DNA is shown in Figure 3. The Rh4 RNA is encoded by 2 exons of similar length (676 and 719 nucleotides), separated by an intron of approximately 9.0 kbase. The 5′ end of the Rh4 RNA, determined by primer extension analysis (M. Fortini and G. Rubin, unpublished observations), is 21 nucleotides upstream of the first nucleotide in the c512-2 cDNA. Thirty nucleotides upstream of the Rh4 transcription initiation site is a sequence similar to the TATA box (Fig. 2), a consensus sequence that frequently precedes the 5′ end of eukaryotic genes transcribed by RNA polymerase II (reviewed in Nevins, 1983). The 3′ end-processing signal AATAAA (reviewed in Platt, 1986) appears in Rh4 twice, overlapping in the 10 nucleotides AATAAATAAA (nucleotides 1355–1364, Fig. 2). Rh3, which shares a 72% identity with Rh4, has a similar 11 nucleotide version of this sequence—AATAAAATAAA—near its 3′ end. The *ninaE*



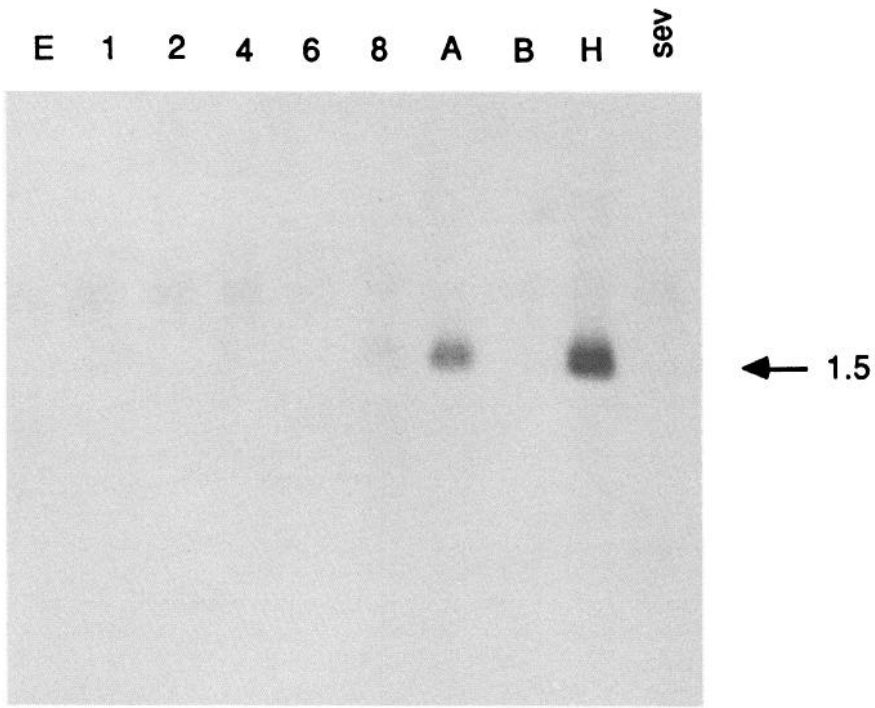


Figure 5. Expression of Rh4 RNA during development and in sevenless (*sev*) adults. Oregon R (P2) strain embryos were collected for 4 hr and incubated at 25°C for 0–16 hr (embryos) and for 1, 2, 4, 6, 8, and 9–10 d (adults). The 1, 2, and 4 d collections coincided approximately with the first, second, and third instar larval stages, and the 6 and 8 d collections corresponded to the early and late pupal periods. Total poly(A)$^+$ RNA was prepared from each of these developmental stages, as well as from separated wild-type and *sev* heads and wild-type bodies. Wild-type head poly(A)$^+$ RNA (0.5 μg) and poly(A)$^+$ RNA for all other samples (2.0 μg) were fractionated on a 3% formaldehyde + 1.5% agarose gel, transferred to nitrocellulose, and probed with a *Pst* I-*Eco* RI c512-2 fragment (nucleotides 1109–1402, Fig. 2) that had been ^{32}P-labeled by nick-translation. Lane *E*, embryos; lanes *1, 2, 4, 6,* and *8,* days of development; *A,* whole adults; *B,* adult bodies; *H,* wild-type heads; and *sev,* sevenless heads. The size of the RNA relative to λ-Hind III markers is given in kilobases. The absence of bands in lanes *E, 1, 2, 4, 6, B,* and *sev* is not due to poor transfer of the RNA to the nitrocellulose, as the expected bands were detected upon reprobing this blot with a *Drosophila* actin DNA probe, pDmA2 (Fyrberg et al., 1983; data not shown).

and Rh2 opsins, which are 67% homologous, both contain a different variation of the 3′ end-processing signal, ATTAAA (see Fig. 4).

Developmental expression of Rh4

Transcription of Rh4 appears to be limited to the last stages of development (Fig. 5). A 1.5 kbase RNA is first detected in RNA prepared from pupae collected on day 8, less than 48 hr prior to eclosion, and at a higher level in the adult, where it is found in the head but not the body. Similar patterns of expression have been observed for the other *Drosophila* rhodopsin genes, as well as for another *Drosophila* gene, *trp* (Montell et al., 1985), which plays a role in phototransduction.

Rh4 expression is localized to photoreceptor cell 7

The spatial localization of Rh4 was determined by *in situ* hybridizations to sections of adult heads and by hybridizing a radioactively labeled Rh4 DNA probe to RNA prepared from the heads of a *Drosophila mutant,* sevenless (*sev*) (Harris et al., 1976) that forms all the photoreceptor cells except for R7. Figure 6 shows the spatial distribution of transcripts homologous to the Rh4-specific DNA probe. The hybridization occurred predominantly to RNA in a subset of cells in the distal region of the compound eye. This is the location of the R7 photoreceptor cells. To confirm that Rh4 is localized exclusively to the R7 cells, RNA prepared from the heads of *sev* adults was hybridized to Rh4 DNA. We found that the 1.5 kbase band present in the heads of wild-type flies is not detected in the lane containing *sev* RNA, even though 4 times as much *sev* RNA was loaded onto the gel (Fig. 5). Thus, within the sensitivity of this analysis, we conclude that Rh4 is expressed specifically in R7 cells. The level of Rh4 expression relative to that of Rh3 was analyzed on an RNA blot using equivalently labeled gene-specific probes. We found that both R7 opsin genes are expressed at similar levels (data not shown).

Rh3 and Rh4 are expressed in nonoverlapping subsets of R7 cells

Rh3 and Rh4 are both expressed in the R7 photoreceptor cells (see above, and Zuker et al., 1987). To determine whether Rh3 and Rh4 are spatially localized to overlapping or nonoverlapping subsets of R7 cells, we performed *in situ* hybridizations to tissue sections. The RNA, in a series of 3 consecutive sections tangential to the surface of the eye, was hybridized with ^{3}H-labeled Rh4 (sections 1 and 3) and with Rh3 (section 2)-specific DNA probes, and the signals were detected by autoradiography. A portion of these sections, as well as a schematic representation of the signal distribution, is shown in Figure 7. Of the R7 cells in the 36 ommatidia included in Figure 7, 21 showed hybridization to the Rh3 probe and 15 to the Rh4 probe. No ommatidium expressed both Rh3 and Rh4. A second series of sections was similarly analyzed to confirm the results depicted in Figure 7 (data not shown). Between the 2 series of sections, we examined a total of 75 ommatidia that could be scored in

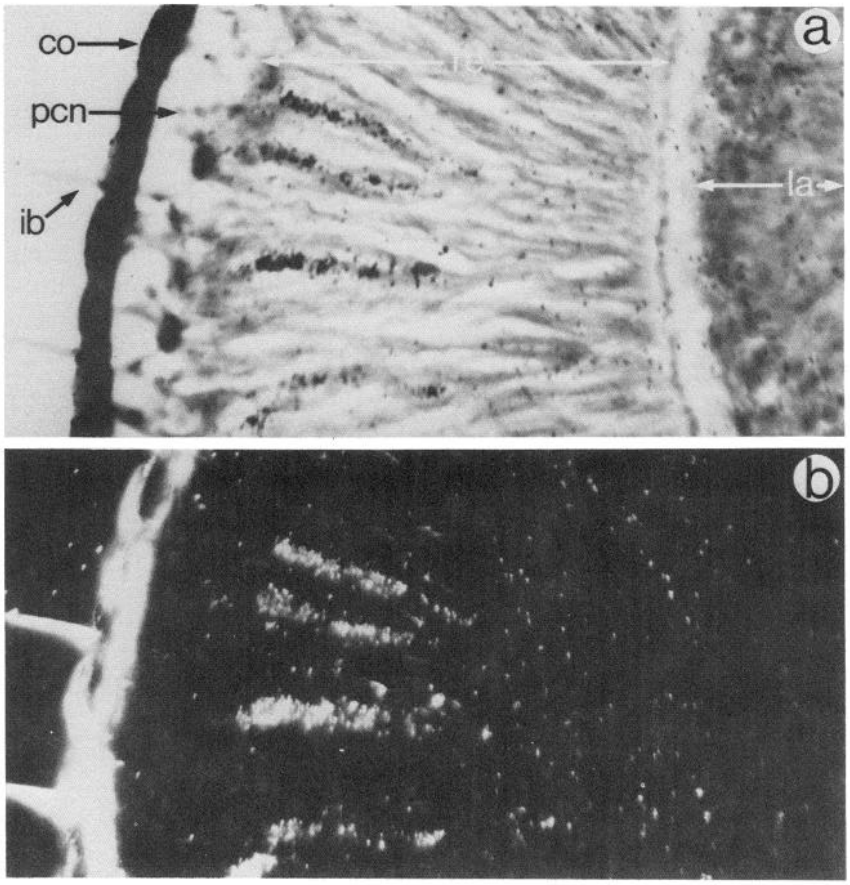


Figure 6. Spatial localization of Rh4 RNA by *in situ* hybridization to horizontal section of an adult fly head. A horizontal section (8 μm) of an adult fly head was prepared and the RNA was hybridized to an Rh4-specific DNA probe nick-translated with ^{3}H-labeled dNTPs. *a,* Bright-field. *b,* Dark-field. *co,* Cornea; *ib,* interommatidial bristles; *la,* lamina ganglionaris; *pcn,* pigment cell nuclei; *re,* retinular cells. In the dark-field image (*b*), some light scattering by the cuticle is seen.

3 consecutive sections. A total of 159 ommatidia were scored in 2 consecutive sections; 78 (49%) showed hybridizations to the Rh3 probe, 81 (51%) to Rh4, and none displayed hybridization signals with both DNA probes. Thus, within the sensitivity of this analysis, it appears that Rh3 and Rh4 are expressed in nonoverlapping subsets of R7 cells. Furthermore, all ommatidia examined expressed either Rh3 or Rh4.

Discussion

Each ommatidium of the fruit fly's compound eye contains 8 photoreceptor cells that fall into 3 classes based on their arrangement and spectral properties (reviewed by Pak and Grabowski, 1978; Hardie, 1983). As a distinct opsin gene is expressed in each photoreceptor cell class, it appears that the differences in spectral properties of these cells are due, at least in part, to the primary structure of the opsin apoprotein. Microspectrophotometric studies (reviewed in Hardie, 1983) have predicted the existence of 2 types of R7 photoreceptor cells that express spectrally distinct photopigments. These 2 R7 types appear to be randomly distributed in the adult eye; the functional significance of this diversity is unknown. In this and the accompanying report (Zuker et al., 1987), we describe 2 opsins, Rh3 and Rh4, that are expressed in nonoverlapping subsets of R7 photoreceptor cells and that may correspond to the previously described spectral types.

In total, 4 *Drosophila* opsin genes have now been isolated. These fall into 2 groups based on amino acid homology comparisons. The *ninaE* (R1–6) and Rh2 (R8) opsins share 67% amino acid identities (Cowman et al., 1986), and the R7 opsins, Rh3 and Rh4, are 72% homologous (Fig. 4A). However, *ninaE* or Rh2 share only about 35% amino acid identity with either R7 opsin.

Although there is considerable amino acid sequence identity between the 4 *Drosophila* opsins, there is little conservation of the number, length, or position of the introns in the genes that encode them (Fig. 8). The introns of the *Drosophila* rhodopsin genes vary in number from 0 to 4, in size from less than 0.2 to approximately 9.0 kbase, and only 1 intron position is conserved among the *Drosophila* visual pigments. The variation in structure among the *Drosophila* opsin genes is surprising when contrasted with the strong conservation of gene structure among vertebrate visual pigments. For example, the human rhodopsin and blue visual pigments share only 42% amino acid identities, while maintaining conservation of all 4 intron positions (Nathans et al., 1986). The only variation in intron–exon structure among the bovine rhodopsin and the 4 human visual pigments is 1 extra intron separating the coding regions of the human green and red pigment genes.

To identify amino acid residues that may be important for rhodopsin function, we compared the sequences of all 4 of the *Drosophila,* 4 human, and 1 bovine visual pigments. Twenty-eight of the 378 (7.4%) amino acid residues in Rh4 are identical among these 9 opsins (Fig. 4B). Twelve of these conserved amino acids are in the 7 membrane-spanning domains. Among these 12 is the lysine (residue 324) in the seventh transmembrane domain, to which the retinal chromophore covalently binds as a Schiff base. Upon absorption of light there is a *cis–trans* isomerization of the retinal that may induce a conformational change in the opsin moeity. Some of these conserved amino acids may be critical for forming the pocket in which the chromophore

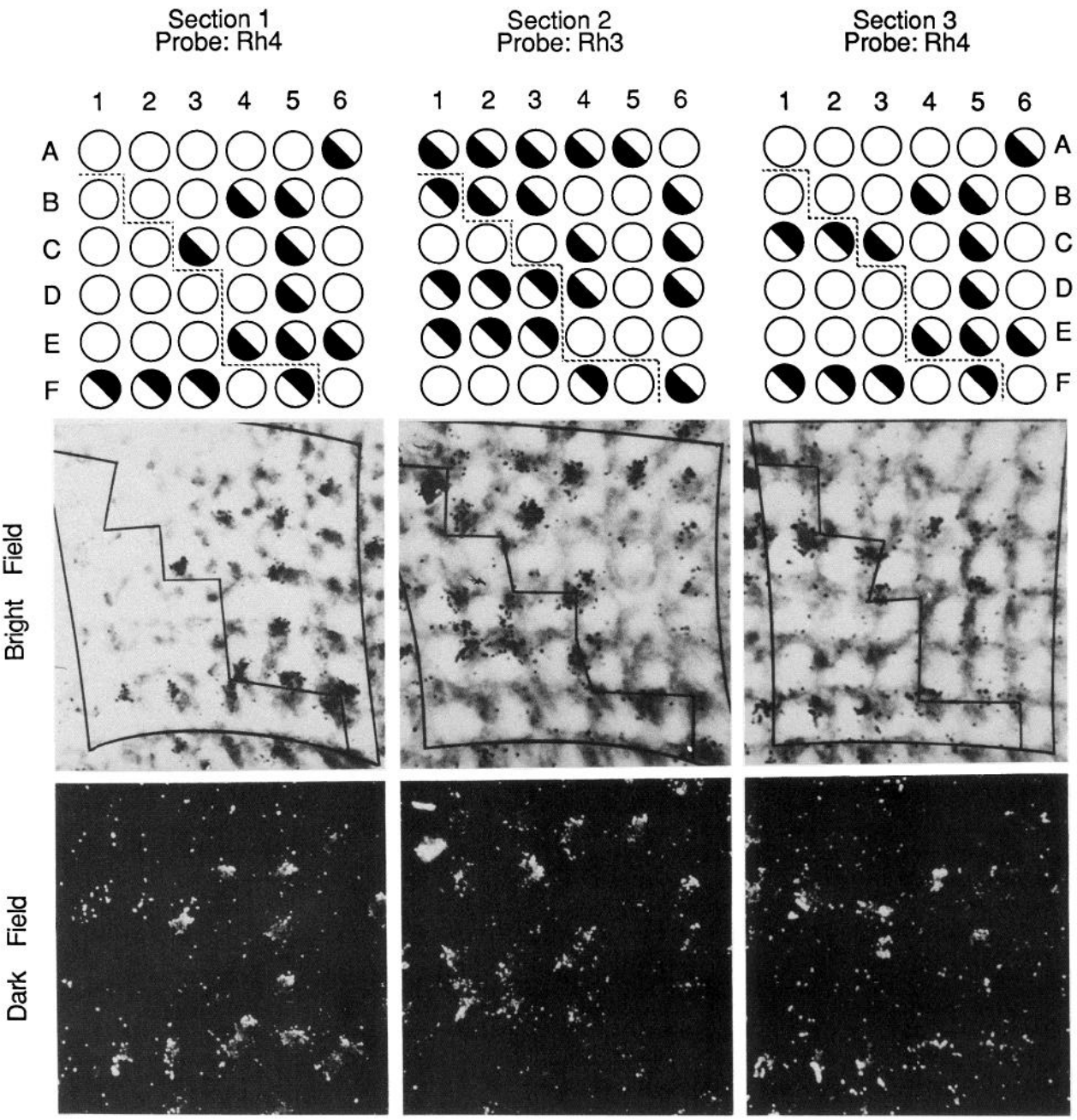


Figure 7. Spatial localization of Rh3 and Rh4 RNA to nonoverlapping subsets of R7 cells. A series of 3 sections (8 μm) tangential to the surface of the eye was prepared, hybridized to Rh4 (*sections 1 and 3*) or Rh3 (*section 2*)-specific DNA probes, which had been nick-translated with ^{3}H-labeled dNTPs, and the hybridization signals were detected by autoradiography. Sections 1–3 were the third, fourth, and fifth sections from the surface of the eye. The photomicrographs of sections 2 and 3 were enlarged relative to section 1 by factors of 1.2 and 1.3, respectively, in order to compensate for the smaller diameter of the ommatidia in the more proximal sections. The *bright-field and dark-field photomicrographs* of each section are shown below the schematic diagrams. In these diagrams, the *circles* represent the centers of the ommatidia and the *shaded* portions indicate hybridization signals. *Dashed lines* traversing diagonally across the schematic diagrams indicate the position of the equator, a plane of mirror-image symmetry that separates the dorsal and ventral portions of the eye. The equator is depicted in each bright-field photomicrograph as a *solid line.* The bright-field photomicrographs are also outlined to indicate the ommatidia considered in the schematic representations. These *outlines* are slightly curved as a consequence of the natural structure of the eye. The dark-field photomicrographs were included to facilitate differentiation between hybridization signals and the dark nuclei in the bright-field images. A positive signal was scored for every cluster of at least 6 silver grains in 1 corner of an ommatidium.. The *signals in* ommatidia *C1* and *C2 of section 1* were not scored, as this portion of section 1 just barely enters the photoreceptor cells. The *region outlined in section 1* is approximately 120 μm wide.

lies or for enabling the conformational change in the opsin to occur after the chromophore isomerizes. It has been proposed that charged (Mathies and Stryer, 1976; Kakitani et al., 1985) and aromatic (Kakitani et al., 1985) amino acids in the transmembrane domains influence the wavelength absorbance of the chromophore. Consistent with this hypothesis, the various opsins differ in their distributions of intramembrane charges and aromatic residues. A negatively charged amino acid is predicted to be in the membrane juxtaposing the protonated retinal Schiff base in all opsins (Oseroff and Callender, 1974). A likely can-

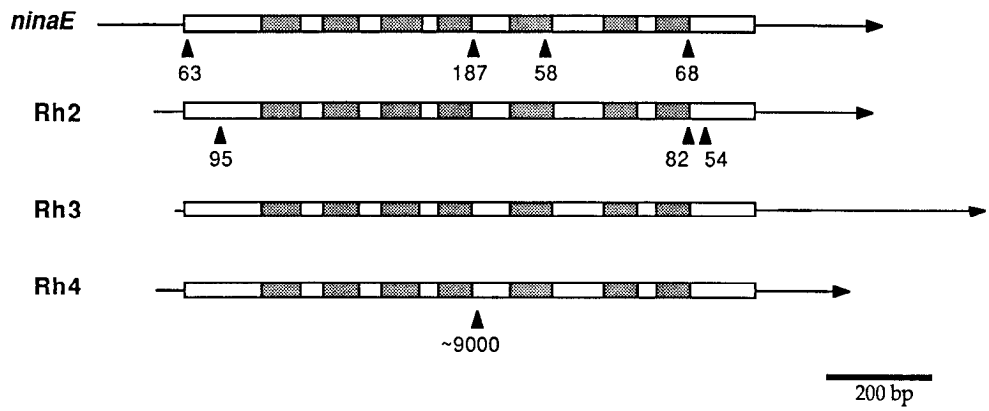


Figure 8. Comparison of intron size, location, and number among the *Drosophila* opsins. The *4 long horizontal lines* represent the RNAs encoding the *Drosophila* opsins: *ninaE,* Rh2, Rh3, and Rh4. The *thick portions of these lines* represent the protein-coding regions. The 5′ and 3′ untranslated regions are depicted by *thin lines. Arrowheads* indicate the direction of transcription. The portions of the thick line shaded with *dots* represent the regions of the RNAs coding for the transmembrane domains. The *unshaded portions* depict the regions coding for cytoplasmic and extracellular protein segments. The lines depicting the opsin RNAs are drawn so that the beginning of the first transmembrane segments are aligned. *Arrowheads perpendicular to the horizontal lines* indicate the positions of the introns in the primary transcripts. Lengths of the introns are shown *below the arrows.* Scale bar represents 200 nucleotides.

didate to provide the counterion is the aspartic acid occupying residue 150 in the third transmembrane domain of Rh4. An acidic amino acid is situated in the corresponding position in all opsin molecules thus far sequenced.

Examination of the amino acid identities along the cytoplasmic face might be useful in identifying important sites of interaction between rhodopsin, the guanosine triphosphate (GTP)-binding protein (Blumenfeld et al., 1985), and other proteins. Eight amino acids are conserved along the cytoplasmic face of all opsins thus far sequenced. Five of these 8 amino acids are in the first cytoplasmic loop. This high degree of conservation strongly suggests that this first cytoplasmic loop is involved in the interaction of rhodopsin with other components of the phototransduction cascade.

The most notable difference between the vertebrate and invertebrate rhodopsins is the presence of an additional 11–12 amino acids in the third cytoplasmic loop of all 4 *Drosophila* rhodopsins. This loop encompasses the region that displays the most extensive conservation among the 4 *Drosophila* opsins; nearly 90% of the first 19 residues in the third cytoplasmic loop are homologous among all 4 *Drosophila* opsins. The functional significance of this region is unclear. However, in contrast to vertebrates, rhodopsin molecules in invertebrates have no rotational or translational mobility (reviewed in Stieve, 1986); it is possible that this domain is involved in the anchoring of rhodopsin to the photoreceptor cell membrane.

Thus far, no specific role has been attributed to the extracellular surface of the rhodopsin molecule. One surprising finding to emerge from the comparison of various opsin genes is that there is as much conservation along the extracellular face of the opsin molecules as along the cytoplasmic face. The second extracellular loop contains the 3 adjacent amino acids, proline, glutamic acid, and glycine (196–198), which are conserved among all opsins except for the glutamic acid in the human red and green cone pigments. Two cysteine residues in the second and third extracellular loops are also fully conserved (126 and 203) and may form a disulfide bond important for the proper conformation of this surface. A glycine in the second and a tryptophan in the third extracellular loops are also invariant. The high conservation of these residues suggests that the extracellular portion of the opsin molecule may have specific roles that have not yet been described.

Finally, on the basis of spectral and physiological criteria, it was predicted that 1 R1–6 opsin, 2 R7-specific opsins, and 1 R8-specific opsin would be expressed in the compound eye of *Drosophila melanogaster* (reviewed in Hardie, 1983). With the identification of a fourth *Drosophila* opsin gene in this report, the minimum number of opsin genes that account for these various classes and subclasses of photoreceptor cells has now been isolated. Through combined application of genetics, *in vitro* mutagenesis, and P-element-mediated germline transformation (Rubin, 1985), it should now be possible to investigate the correlation between the structure and function of each of these opsins.

References

Bankier, A. T., and B. G. Barrell (1983) Shotgun DNA sequencing. In *Techniques in Nucleic Acid Biochemistry,* vol. B5, R. A. Flavell, ed., pp. 1–33, Elsevier, Amsterdam.

Blumenfeld, A., J. Erusalimsky, O. Heichal, Z. Selinger, and B. Minke (1985) Light-activated guanosinetriphosphatase in *Musca* eye membranes resembles the prolonged depolarizing afterpotential in photoreceptor cells. Proc. Natl. Acad. Sci. USA *82:* 7116–7120.

Cowman, A. F., C. S. Zuker, and G. M. Rubin (1986) An opsin gene expressed in only one photoreceptor cell type of the *Drosophila* eye. Cell *44:* 705–710.

Dente, L., G. Cesareni, and R. Cortese (1983) pEMBL: A new family of single stranded plasmids. Nucleic Acids Res. *11:* 1645–1655.

Frischauf, A.-M., H. Lehrach, A. Poustka, and N. Murray (1983) Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. *170:* 827–842.

Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson (1983) Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. Cell *33:* 115–123.

Hafen, E., M. Levine, R. L. Garber, and W. J. Gehring (1983) An improved *in situ* hybridization method for detection of cellular RNAs in *Drosophila* tissue sections and its application for localizing transcripts of the homeotic *Antennapedia* gene complex. EMBO J. *2:* 617–623.

Hardie, R. C. (1983) Functional organization of the fly retina. In *Progress in Sensory Physiology,* H. Autrum, D. Ottoson, E. R. Perl, R. F. Schmidt, H. Shimazu, and W. D. Willis, eds., pp. 1–81, Springer, New York.

Hargrave, P. A., J. H. McDowell, D. R. Curtis, J. K. Wang, E. Juszczak, S.-L. Fong, J. K. Mohanna Rao, and P. Argos (1983) The structure

of bovine rhodopsin. Biophys. Struct. Mech. *9:* 235–244.
Harris, W. A., W. S. Stark, and J. A. Walker (1976) Genetic dissection of the photoreceptor system in the compound eye of *Drosophila melanogaster.* J. Physiol. (Lond.) *256:* 415–439.
Kakitani, H., T. Kakitani, H. Rodman, and B. Honig (1985) On the mechanism of wavelength regulation in visual pigments. Photochem. Photobiol. *41:* 471–479.
Kyte, J., and R. F. Doolittle (1982) A simple model for displaying the hydropathic character of a protein. J. Mol. Biol. *157:* 105–132.
Levy, L. S., R. Ganguly, N. Ganguly, and J. E. Manning (1982) The selection, expression, and organization of a set of head-specific genes in *Drosophila.* Dev. Biol. *94:* 451–464.
Maniatis, T., E. F. Fritsch, and J. Sambrook (1982) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Mathies, R., and L. Stryer (1976) Retinal has a highly dipolar vertically excited singlet state: Implications for vision. Proc. Natl. Acad. Sci. USA *73:* 2169–2173.
Montell, C., K. Jones, E. Hafen, and G. Rubin (1985) Rescue of the *Drosophila* phototransduction mutation *trp* by germline transformation. Science *230:* 1040–1043.
Nathans, J., and D. S. Hogness (1983) Isolation, sequence analysis, and intron–exon arrangement of the gene encoding bovine rhodopsin. Cell *34:* 807–814.
Nathans, J., and D. S. Hogness (1984) Isolation and nucleotide sequence of the gene encoding human rhodopsin.. Proc. Natl. Acad. Sci. USA *81:* 4851–4855.
Nathans, J., D. Thomas, and D. S. Hogness (1986) Molecular genetics of human color vision: The genes encoding blue, green, and red pigments. Science *232:* 193–202.
Nevins, J. R. (1983) The pathway of eukaryotic mRNA formation. Annu. Rev. Biochem. *52:* 441–466.
Oseroff, A. R., and R. H. Callender (1974) Resonance raman spectroscopy of rhodopsin in retinal disk membranes. Biochemistry *13:* 4243–4248.
Ostroy, S. E., M. Wilson, and W. L. Pak (1974) *Drosophila* rhodopsin: Photochemistry, extraction and differences in the *NorpA*^{P12} phototransduction mutant. Biochem. Biophys. Res. Commun. *59:* 960–966.
O'Tousa, J. E., W. Baehr, R. L. Martin, J. Hirsch, W. L. Pak, and M. L. Applebury (1985) The *Drosophila ninaE* gene encodes an opsin. Cell *40:* 839–850.
Ovchinnikov, Y. A. (1982) Rhodopsin and bacteriorhodopsin: Structure–function relationships. FEBS Lett. *148:* 179–189.
Pak, W. L., and S. R. Grabowski (1978) Physiology of the visual and flight systems. In *The Genetics and Biology of Drosophila,* vol. 2a, M. Ashburner and T. R. F. Wright, eds., pp. 553–604, Academic, New York.
Platt, T. (1986) Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. *55:* 339–372.
Rubin, G. M. (1985) P transposable elements and their use as genetic tools in *Drosophila.* Trends Neurosci. *8:* 231–233.
Sanger, F., S. Nicklen, and A.R. Coulson (1977) DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA *74:* 5463–5467.
Stieve, H. (1986) Introduction. In *The Molecular Mechanism of Photoreception,* H. Stieve, ed., pp. 1–10, Springer, Berlin.
Zuker, C. S., A. F. Cowman, and G. M. Rubin (1985) Isolation and structure of a rhodopsin gene from *D. melanogaster.* Cell *40:* 851–858.
Zuker, C. S., C. Montell, K. Jones, T. Laverty, and G. M. Rubin (1987) A rhodopsin gene expressed in photoreceptor cell R7 of the *Drosophila* eye: Homologies with other signal-transducing molecules. J. Neurosci. *7:* 1537–1550.