Variable Rates of Evolution Among Drosophila Opsin Genes

John P. Carulli and Daniel L. Hartl

Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110

Manuscript received January 17, 1992
Accepted for publication May 29, 1992

ABSTRACT

DNA sequences and chromosomal locations of four *Drosophila pseudoobscura* opsin genes were compared with those from *Drosophila melanogaster*, to determine factors that influence the evolution of multigene families. Although the opsin proteins perform the same primary functions, the comparisons reveal a wide range of evolutionary rates. Amino acid identities for the opsins range from 90% for Rh2 to more than 95% for Rh1 and Rh4. Variation in the rate of synonymous site substitution is especially striking: the major opsin, encoded by the *Rh1* locus, differs at only 26.1% of synonymous sites between *D. pseudoobscura* and *D. melanogaster*, while the other opsin loci differ by as much as 39.2% at synonymous sites. *Rh3* and *Rh4* have similar levels of synonymous nucleotide substitution but significantly different amounts of amino acid replacement. This decoupling of nucleotide substitution and amino acid replacement suggests that different selective pressures are acting on these similar genes. There is significant heterogeneity in base composition and codon usage bias among the opsin genes in both species, but there are no consistent relationships between these factors and the rate of evolution of the opsins. In addition to exhibiting variation in evolutionary rates, the opsin loci in these species reveal rearrangements of chromosome elements.

THE opsin gene family provides an excellent opportunity to study gene duplication and divergence in the evolution of new biological functions. Opsins form the protein component of visual pigments, and they are ubiquitous in the animal kingdom (MARTIN *et al.* 1986). Physiological studies suggest that many animals have multiple visual pigments (MENZEL 1979), and molecular cloning of the human (NATHANS and HOGNESS 1984; NATHANS *et al.* 1986; NATHANS 1987; YOKOYAMA and YOKOYAMA 1989) and Drosophila (O'TOUSA *et al.* 1985; ZUKER, COWMAN and RUBIN 1985; COWMAN, ZUKER and RUBIN 1986; FRYXELL and MEYEROWITZ 1987; MONTELL *et al.* 1987; ZUKER *et al.* 1987) opsins has shown that they are encoded by multigene families. In both humans and flies, opsin genes have duplicated, dispersed to different genomic locations, acquired cell-specific expression patterns, and evolved the capacity to respond optimally to different wavelengths of light. Phylogenetic analyses indicate that the duplication of opsin genes occurred independently in the lines of descent leading to mammals and insects (FRYXELL and MEYEROWITZ 1991; ZUKER *et al.* 1987).

Four genes encoding opsin proteins have been identified in *Drosophila melanogaster* (O'TOUSA *et al.* 1985; ZUKER, COWMAN and RUBIN 1985; COWMAN, ZUKER and RUBIN 1986; FRYXELL and MEYEROWITZ 1987; MONTELL *et al.* 1987; ZUKER *et al.* 1987). While the individual rhodopsins (opsin protein plus retinal chromophore) perform the same primary functions, each has a unique role in the visual system of the fly (HARRIS, STARK and WALKER 1976). Each of the 700–800 ommatidia in the compound eyes has eight rhodopsin-containing photoreceptor cells, R1 through R8, arranged in a stereotypical pattern (PAK and GRABOWSKI 1978). The major rhodopsin, encoded by the *ninaE* (*Rh1*) locus, is expressed in photoreceptor cells R1–6 (O'TOUSA *et al.* 1985; POLLOCK and BENZER 1988; COWMAN, ZUKER and RUBIN 1986), and is optimally sensitive to 480 nm light (*e.g.*, HARRIS, STARK and WALKER 1976; O'TOUSA *et al.* 1985; COWMAN, ZUKER and RUBIN 1986). Two ultraviolet (370 nm)-sensitive rhodopsins, Rh3 and Rh4, are expressed in nonoverlapping sets of R7 photoreceptor cells in the ommatidia (MONTELL *et al.* 1987; ZUKER *et al.* 1987; FORTINI and RUBIN 1990). The gene for the R8 opsin has not been identified (FORTINI and RUBIN 1990). The three dorsally located simple eyes, the ocelli, express the Rh2 opsin protein (COWMAN, ZUKER and RUBIN 1986, POLLOCK and BENZER 1988), which is sensitive to 450-nm light (HARRIS, STARK and WALKER 1976; HU, REICHERT and STARK 1978). The four opsin genes can be divided into two groups based on amino acid identity: Rh1 and Rh2 are 67% identical with one another at the amino acid level; Rh3 and Rh4 are 75% identical at the amino acid level. The four opsins together share 35% amino acid identity (ZUKER *et al.* 1987).

We have cloned and sequenced *Rh1*, *Rh2*, *Rh3* and *Rh4* from *Drosophila pseudoobscura*, a species in the *obscura* group in the subgenus *Sophophora* (LAKOVAARA and SAURA 1982). The lineages leading to the *obscura* group and the *melanogaster* group diverged approximately 30 to 40 million years ago (BEVERLY

and WILSON 1984). The DNA sequences of several genes from these two species have been compared, and the rate of nucleotide substitution among genes has been shown to vary considerably (RILEY 1988). The comparisons that have been made, however, involve a wide variety of genes which are likely to be subject to widely different functional constraints. The opsins are an ideal system for studying amino acid and nucleotide sequence evolution within a family of genes that appear to have similar functional constraints. In this paper, DNA sequences and inferred amino acid sequences of the *D. pseudoobscura* opsins are aligned with their *D. melanogaster* homologs, and patterns of nucleotide substitution and amino acid replacement are determined. Amino acid replacements in the opsins are examined with respect to the known structural and functional properties of Drosophila and vertebrate opsins (O'TOUSA *et al.* 1985; NATHANS 1987; ZUKER *et al.* 1987). Rates and patterns of nucleotide substitution are compared among the different opsin loci, and two factors that are thought to affect rates of synonymous nucleotide substitution are analyzed: codon usage bias, and variation in base composition within the genome (LI and GRAUR 1991; WOLFE, SHARP and LI 1989; SHIELDS *et al.* 1988).

MATERIALS AND METHODS

Library screening: A *Drosophila pseudoobscura* genomic library, constructed from *Mbo*I partial digests inserted into the *Bam*HI site of EMBL3, was provided by STEPHEN W. SCHAEFFER (SCHAEFFER and AQUADRO 1987). The DNA probes used to identify *Rh1*, *Rh3*, the 5′ exon of *Rh4* and *seven in absentia* (*sina*) clones were derived from plasmid clones of *D. melanogaster* opsins (provided by JOSEPH O'TOUSA) or *sina* (provided by RICHARD CARTHEW). Probes from *D. melanogaster* and *D. pseudoobscura* genomic DNA for *Rh2* and the 5′ exon of *Rh4* were obtained using the polymerase chain reaction with the following primer pairs:

Rh2 A, 5′ CATTTGCCGGAGACGCCATTCGA 3′;
Rh2 B, 5′ GCCTTTGAATCCGCCTCAGATGT 3′;
Rh4 A, 5′ ATGGAGCCGTTCTCGAACGC 3′;
Rh4 B, 5′ CTGGCACGAATCGATCCC 3′.

Probes for DNA hybridizations were radioactively labeled by random priming and incorporation of [α-^{32}P]dATP with DNA polymerase (FEINBERG and VOGELSTEIN 1983).

Subcloning: DNA from EMBL3 clones carrying opsin genes was prepared by the method of HELMS *et al.* (1985). DNA from these clones and from *D. pseudoobscura* was digested with several restriction enzymes, either singly or in pairs, and the restriction fragments were resolved on 0.8% agarose gels. The DNA was transferred to nylon membranes by standard methods (SAMBROOK, FRITSCH and MANIATIS 1989) or by using a Posi-blot pressure blotter (Stratagene). Restriction fragments encompassing the opsin genes were identified by hybridization with radioactive probes. The appropriate restriction fragments were subcloned into pBluescript (Stratagene) and transformed into XL1-blue by electroporation (Bio-Rad). The clones used in this study include: two *Rh1* clones: a 1.2-kb *Eco*RI fragment and a 5.5-kb *Hin*dIII fragment; *Rh2*: 4.0-kb *Eco*RI fragment; two *Rh3* clones: a 900-bp *Cla*I fragment and a 6.5-kb *Bam*HI-*Sal*I fragment; 5′ exon *Rh4*: 4.4-kb *Eco*RI fragment; *sina*: 6.3-kb *Eco*RI fragment.

***In situ* hybridizations:** Polytene chromosomes were hybridized with biotinylated probes as described in AJIOKA *et al.* (1991).

DNA sequencing: Double-stranded plasmid DNAs prepared by alkaline lysis (SAMBROOK, FRITSCH and MANIATIS 1989) were used as templates for most of the DNA sequencing. Plasmid DNA was denatured using 0.2 M NaOH and precipitated with ethanol prior to sequencing with Sequenase (USB) according to manufacturer's recommendations. The 3′ exon of *Rh4* was sequenced directly from an EMBL3 clone (DNA prepared as cited above) using *Taq* polymerase in the thermal cycling method of CRAXTON (1991). Initial sequencing reactions utilized universal primers or oligonucleotides complementary to selected sites in the published *D. melanogaster* opsin sequences. Additional oligonucleotide primers were prepared as needed to sequence both strands of each of the *D. pseudoobscura* opsin genes.

Sequence analysis: Many of the sequence analyses were carried out using the GCG package (DEVEREUX, HAEBERLI and SMITHIES 1984). DNA sequences of the *D. pseudoobscura* sequences presented in this paper were aligned with published sequences for the *D. melanogaster* opsins (GENBANK accession numbers KO2315, M12896, M17718, M17719 and M17730) using the GAP program. Coding sequences of the *D. pseudoobscura* opsins were inferred from alignments with the *D. melanogaster* homologs. Protein sequences were obtained using the TRANSLATE function from GCG and aligned using GAP. Nucleotide substitution rates in the coding regions, including estimates of multiple substitutions, were calculated using the SYNSUB program, provided by RICHARD LEWONTIN (LEWONTIN 1988). The equilibrium codon frequencies for the substitution matrices, which are used in the estimates of "evolutionary events" (multiple substitutions) per codon under the LEWONTIN (1988) model, were determined using two methods: the mean codon usage of the two sequences being compared; and the mean codon usage over all eight of the opsin sequences.

Two different measures of codon bias were calculated: the frequency of optimal codons (F_{op}) (IKEMURA 1985) and χ^2/L (SHIELDS *et al.* 1988). F_{op} is equal to the number of optimal codons in a coding sequence divided by the total number of codons (minus those coding for methionine and tryptophan). Optimal codons are defined as those codons that are used with a statistically significant higher frequency in a reference set of "highly expressed" genes relative to a reference set of "lowly expressed" genes (P. SHARP, personal communication). F_{op} values for each of the opsin genes in both species were calculated using the codons determined to be optimal in *D. melanogaster* (P. SHARP, personal communication). χ^2/L is the chi-square for deviation from equal usage of all codons within a synonymous group, divided by the total number of codons (minus those coding for methionine and trytophan).

RESULTS

Chromosomal locations of opsin genes in *D. pseudoobscura*: *In situ* hybridization of biotinylated DNA probes to polytene chromosomes of *D. pseudoobscura* was used to determine the chromosomal locations of the opsin genes. Figure 1 shows that *Rh1* (section 45), *Rh2* (section 52) and *Rh3* (section 53) are all on the second chromosome of *D. pseudoobscura*, whereas *Rh4* (section 38) is on the right arm of the *X* chromosome.

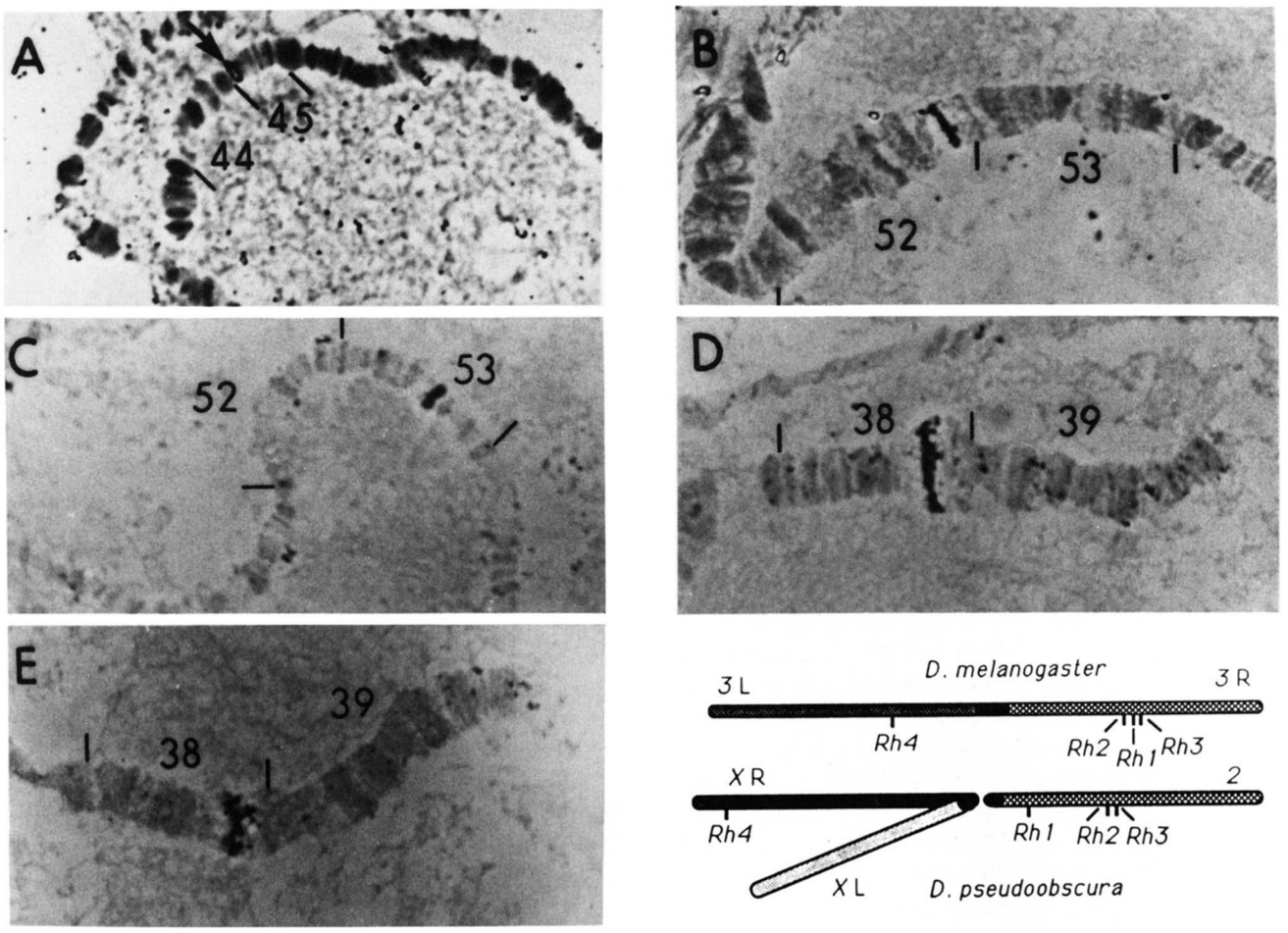

FIGURE 1.—*In situ* hybridizations of *D. pseudoobscura* opsin clones to polytene chromosomes. (A) An *Rh1* plasmid clone, Dprh1.2, hybridizes to polytene section 45 in *D. pseudoobscura* chromosome 2. (B) A lambda clone, λ115, which encompasses the *Rh2* gene, hybridizes to polytene section 52 on *D. pseudoobscura* chromosome 2. (C) λ128, which contains the coding sequences for *D. pseudoobscura Rh3*, hybridizes to polytene section 53, also on chromosome 2. (D) A lambda clone, sina-2, which has the sequences for the 5′exon of *Rh4* and for most of *seven in absentia* is localized to polytene section 38 on the right arm of the *D. pseudoobscura X* chromosome. (E) The 3′exon of *Rh4* is on another lambda clone, λ4309, that also hybridizes to polytene section 38, at a site closely adjacent to sina-2. (F) A diagram (not to scale) showing the relative positions of the opsin genes on *D. melanogaster* and *D. pseudoobscura* chromosomes. Homologous chromosome arms have the same shading, centromeres are black. *D. melanogaster 3L* is homologous to *D. pseudoobscura XR*, and *D. melanogaster* is *3R* is homologous to *D. pseudoobscura* chromosome 2.

In *D. melanogaster* the opsin genes also fall into two groups: *Rh1* (section 92B6-7), *Rh2* (91D1-2) and *Rh3* (92D) on *3R*, and *Rh4* (73D3-5) on *3L* (COWMAN, ZUKER and RUBIN 1986; FRYXELL and MEYEROWITZ 1987; MONTELL *et al.* 1987; O'TOUSA *et al.* 1985; ZUKER, COWMAN and RUBIN 1985; ZUKER *et al.* 1987). These observations are consistent with genetic and molecular cytological data that have established that chromosome *2* of *D. pseudoobscura* is homologous to *3R* of *D. melanogaster* and chromsome *XR* of *D. pseudoobscura* is homologous to *3L* of *D. melanogaster* (ASHBURNER 1989; LAKOVAARA and SAURA 1982; STEINEMANN, PINSKER and SPERLICH 1984). Although the opsin genes are on homologous chromosome arms in these two species, the positions of the genes relative to one another and to the telomeres and centromeres of their respective chromosomes are different. In *D. melanogaster*, *Rh1*, *Rh2* and *Rh3* are in adjacent polytene sections, with *Rh1* located between *Rh2* and *Rh3*. In *D. pseudoobscura*, *Rh1* is separated from *Rh2* and *Rh3* by several sections. *Rh4* is near the middle of *3L* in *D. melanogaster*, and is a few polytene sections away from the telomere on *XR* in *D. pseudoobscura*.

Molecular organization of the *D. pseudoobscura* opsin genes: The organization of the *D. pseudoobscura* opsin genes was inferred by alignment of the DNA sequences with the published sequences of the *D. melanogaster* homologs. The major structural features of the genes, including initiation codons, introns, and stop codons were readily apparent from the alignments. There are no insertions or deletions more than two amino acids in length in any of the opsins. The intron positions in the *D. pseudoobscura* opsin genes are identical to those in *D. melanogaster*. The total length of intron sequences in *Rh1* is much greater in *D. pseudoobscura*, primarily because the second intron

in this species is 376 bp long, *vs.* 188 nucleotides in *D. melanogaster*. Total intron length in *Rh2* is similar in the two species. *Rh3* does not contain an intron in either species. Interestingly, *Rh4* has a large intron with another gene, *seven in absentia* (*sina*), nested within it in *D. melanogaster* (CARTHEW and RUBIN 1990), while in *D. virilis Rh4* has no intron and *sina* is located elsewhere on the same chromosome arm (NEUFELD, CARTHEW and RUBIN 1991). In *D. pseudoobscura*, the nested arrangement appears to be the case, since the 5′ exon of *Rh4* is on the same lambda clone as *sina* and a separate lambda clone containing the 3′ exon hybridizes to an adjacent chromosomal site.

DNA sequences of the *D. pseudoobscura* opsin genes: The DNA sequences of the opsin genes are shown in Figure 2. The coding regions, introns, and flanking regions were aligned independently for each comparison between the *D. pseudoobscura* and *D. melanogaster* genes. The coding regions of the *D. melanogaster* opsins are shown above the *D. pseudoobscura* sequences. Alignment of sequences 5′ to the coding region shows that several well characterized *cis*-acting regulatory regions, identified because of their similarity in *D. melanogaster* and *D. virilis* (FORTINI and RUBIN 1990), are also conserved in all of the *D. pseudoobscura* opsins. Introns from homologous positions in *Rh1* and *Rh2* were also aligned. With the exception of splice donor and acceptor splice sites, there is little sequence conservation within the introns, and in no case was more than 50% identity between the two species observed (alignments not shown).

Analysis of nucleotide substitutions: The amount of nucleotide substitution in the coding regions (Table 1) varies significantly among the opsin genes ($\chi^2 = 17.26$, $P < 0.01$). Nonsynonymous substitutions, which result in amino acid replacements, range from 3.3% for *Rh1* to 9.3% for *Rh2*. Synonymous site differences at the *Rh1* locus are 26.1%, which is among the lower values reported for genes that have been compared between these species (RILEY 1988; SHARP and LI 1989; SHIELDS *et al.* 1988). By contrast, the synonymous site difference of 39.2% for the *Rh2* comparison is among the higher values reported for comparisons of DNA sequences between these two species (RILEY 1988; SHARP and LI 1989; SHIELDS *et al.* 1988). The lower level of synonymous substitution in *Rh1* is not associated with the amount of amino acid replacement: *Rh1* and *Rh4* have similar amounts of amino acid replacement, but the amount of synonymous site substitution in these two genes differs significantly.

The number of evolutionary substitutions per synonymous codon in the opsin genes is shown in Table 1. The choice of LEWONTIN's (1988) method to estimate multiple substitutions was motivated by the fact that this method analyzes the data at the level of the codon rather than at the level of the nucleotide. Constraints on amino acid replacements in most protein coding sequences strongly influence the probabilities of nucleotide substitution, so that substitution (divergence) estimates based solely on analysis of nucleotides can be misleading (LEWONTIN 1988). For each comparison, similar estimates of the number of evolutionary events were obtained using two different codon usage tables for the SYNSUB program: the average codon usage of the two genes being compared (Table 1); and the average codon usage in all eight of the Drosophila opsin sequences analyzed in this paper (data not shown). Our data are consistent with the analyses of RILEY (1988), in that different degeneracy groups (*i.e.*, 2-, 3-, 4- and 6-fold degenerate codons) evolve at different rates in some of the opsin genes.

Base composition of the opsin loci: To examine the possibility that the differences in evolutionary rate among the opsin genes are due to variable mutation rates in different parts of the genome, we compared the base composition of noncoding DNA (flanking regions and introns) of each locus with the base composition of third base positions (the location of most synonymous sites). This approach has been used in the analysis of the structuring within mammalian genomes, and has revealed mutational bias at silent sites toward a particular G + C content (WOLFE, SHARP and LI 1989). Regional variation in base composition is also known to have an effect on the evolutionary rates of mammalian genes (LI and GRAUR 1991), but the effect of such variation in Drosophila has only recently been addressed (MORIYAMA and GOJOBORI 1992).

Table 2 shows that significant variation in base composition is observed among opsin loci in both *D. pseudoobscura* ($\chi^2 = 32.7$, $P < 0.01$) and *D. melanogaster* ($\chi^2 = 19.1$, $P < 0.01$), but there is only a weak correlation ($r = 0.557$, $P > 0.1$) between the G + C content of the noncoding DNA and the G + C content of silent sites within the eight genes. Tests for correlation between the G + C content and rate of evolution of the opsin loci (Table 3) show that there are no consistent relationships between base composition and rate of evolution in this sample. It is noteworthy, however, that homologous loci in *D. pseudoobscura* and *D. melanogaster* can differ significantly in G + C content. *Rh2* shows the most dramatic difference: the entire region is 53.3% G + C in *D. pseudoobscura vs.* 44.2% G + C in *D. melanogaster* ($\chi^2 = 28.6$, $P < 0.01$), and the noncoding DNA is 45.7% G + C in *D. pseudoobscura vs.* 33.1% in *D. melanogaster* ($\chi^2 = 26.7$, $P < 0.01$).

Codon usage bias varies significantly among the opsin loci: Total codon usage does not vary significantly among the opsins in either *D. melanogaster* or *D. pseudoobscura*. Chi-square tests of the heterogeneity of codon usage among all four genes (including only

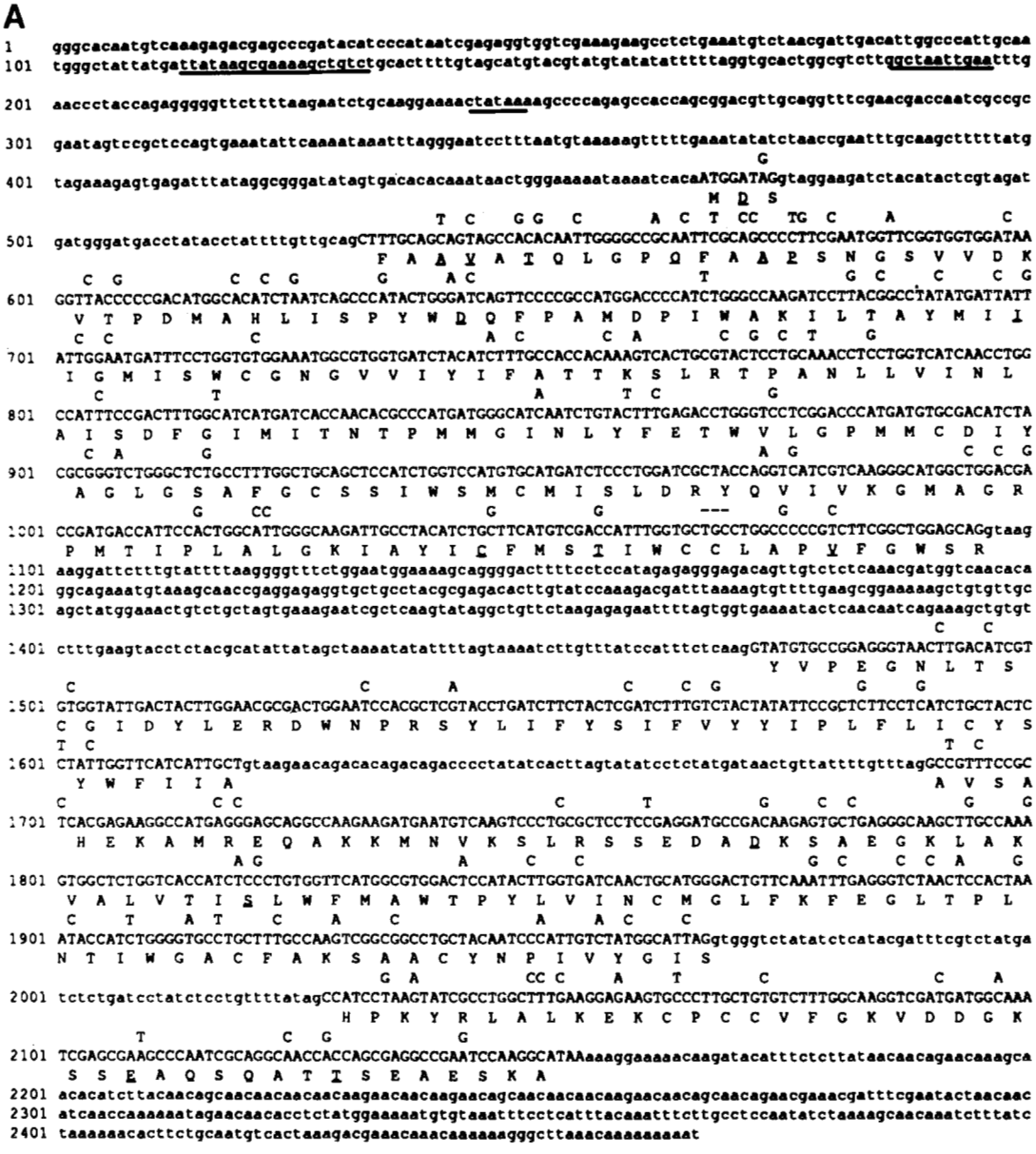

FIGURE 2.—The DNA sequence of *D. pseudoobscura Rh1*, *Rh2*, *Rh3* and *Rh4*. The coding sequences are in uppercase letters, with flanking regions (including regions that are presumably transcribed) and introns in lower case. The single letter amino acid code is printed beneath the second nucleotide of each codon. Amino acids that differ between *D. pseudoobscura* and *D. melanogaster* are underlined. Differences between the *D. melanogaster* and *D. pseudoobscura* coding regions are shown above the *D. pseudoobscura* sequence. *D. melanogaster* DNA sequences outside of the coding region are not shown. *cis*-Acting regulatory domains in the 5′ flanking DNA, previously identified by FORTINI and RUBIN (1990), are underlined. Most of these domains are highly conserved among *D. pseudoobscura*, *D. melanogaster* and *Drosophila virilis*. (A) The *Rh1* (*ninaE*) DNA sequence. The positions of the introns are conserved in *D. pseudoobscura* and *D. melanogaster*. The sizes of the introns in both species are as follows: *D. pseudoobscura* (D.p.) intron 1 = 63 bp, *D. melanogaster* (D.m.) intron 1 = 63 bp; D.p. intron 2 = 376 bp, D.m. intron 2 = 187 bp; D.p. intron 3 = 58 bp, D.m. intron 3 = 70 bp; D.p. intron 4 = 68 bp, D.m. intron 4 = 60 bp. (B) The *Rh2* DNA sequence. The intron positions in *Rh2* are also conserved between *D. pseudoobscura* and *D. melanogaster*. The sizes of the introns are as follows: D.p. intron 1 = 67 bp, D.m. intron 1 = 96 bp; D.p. intron 2 = 64 bp, D.m. intron 2 = 82 bp; D.p. intron 3 = 63 bp, D.m. intron 3 = 56 bp. (C) The *Rh3* DNA sequence. There are no introns in *Rh3* in *D. pseudoobscura* or *D. melanogaster*. (D) The *Rh4* DNA sequence. The position of the large intron, which is conserved in *D. pseudoobscura* and *D. melanogaster*, is marked by a dotted line. Additional DNA sequences from the intron, adjacent to both of the exons, were determined but are not shown. (Figure 2 continued on pages 198–200.)

those codons that are used 20 times or more in the four genes combined) do not reveal significant variability: for *D. melanogaster*, $\chi^2 = 65.61$, d.f. = 60, $P > 0.1$; for *D. pseudoobscura*, $\chi^2 = 77.89$, d.f. = 57, $P > 0.1$. Despite similarities in general codon usage among the opsins, there are considerable differences in codon usage bias among the genes (Table 2). The total number of "optimal" codons used in each gene varies significantly among the opsin genes within both species (Table 2): for *D. pseudoobscura* $\chi^2 = 7.52$, d.f. = 3, $0.05 < P < 0.1$; for *D. melanogaster* $\chi^2 = 29.3$, d.f. = 3, $P < 0.01$. χ^2/L values also span a wide range

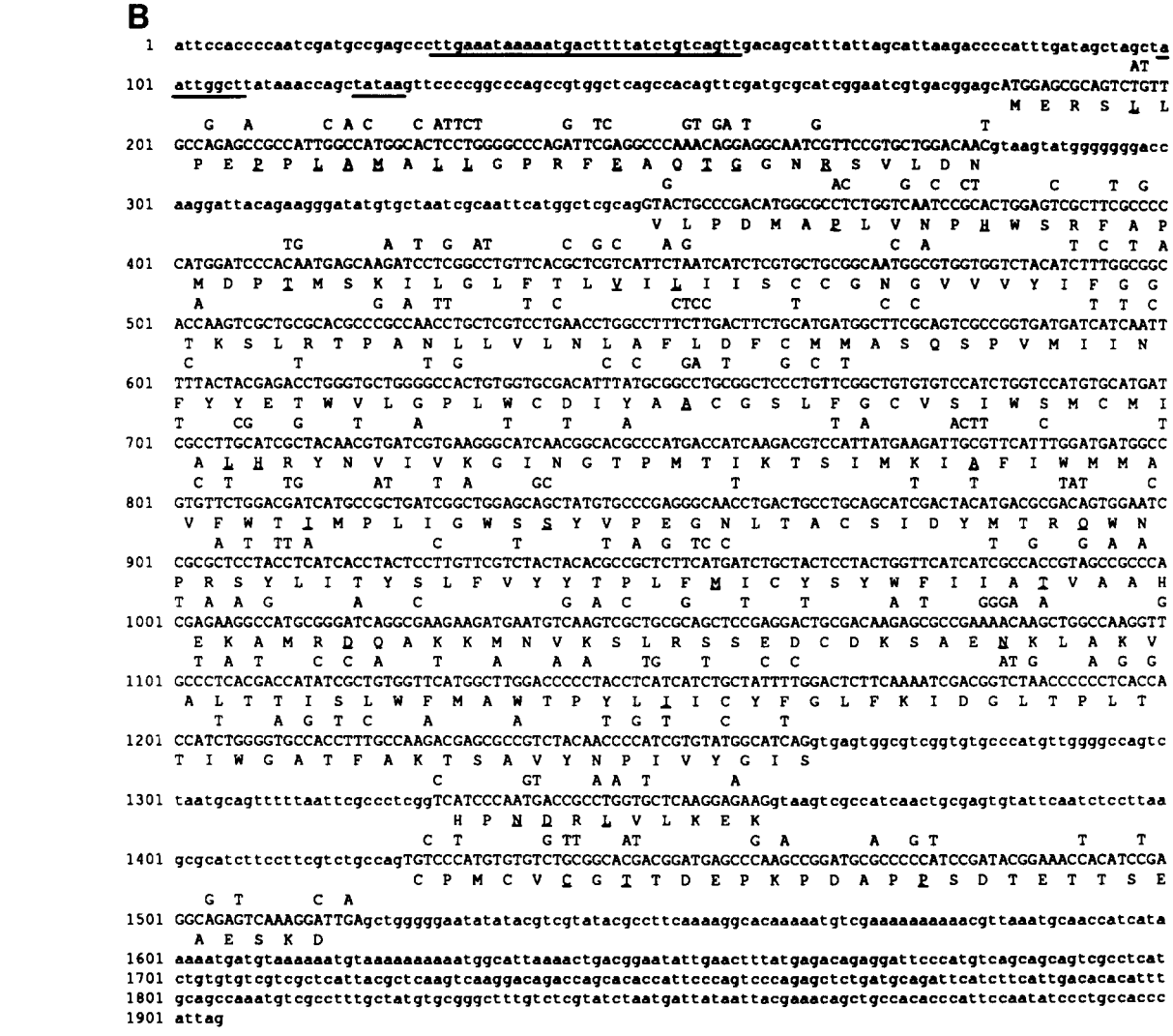

FIGURE 2.—Continued

similar to that of the F_{op} values. We include χ^2/L because it does not require *a priori* decisions regarding optimal codons, and is therefore indifferent to potential differences in codon preference in the two species. Although the numbers for F_{op} and χ^2/L are not directly comparable because of fundamental differences in computation, the two measures are significantly correlated ($r = 0.755$, d.f. = 6, $P < 0.05$). There is no consistent association between codon usage bias and the rate of evolution of the opsins. Synonymous site substitution shows a significant negative correlation with the *D. melanogaster* codon bias measures, but no such correlation exists with the *D. pseudoobscura* values (Table 3).

Comparison of the amino acid sequences of the opsins: Amino acid differences between each of the *D. pseudoobscura* opsins and its *D. melanogaster* homolog are summarized in Figure 3 and Table 1. Each of the *D. pseudoobscura* opsin proteins is at least 90% identical at the amino acid level with its *D. melanogaster* homolog. The range of identity values (Table 1) varies significantly ($\chi^2 = 15.5$, $P < 0.01$), but is within the range of 85.5–96.5% reported for other genes that have been compared between the two species (RILEY 1988).

In Figure 3, the amino acid comparisons are summarized in a format that facilitates comparison of the structural domains of the rhodopsins. Each opsin has seven hydrophobic transmembrane domains connected by loops of polar amino acids that extend into the cytoplasm or extracellular matrix. The chromophore-bearing lysine in the seventh transmembrane domain is conserved in all four *D. pseudoobscura* opsins. Rates of amino acid replacement vary significantly among the different structural and functional domains of the proteins. Only 27.8% of the amino acid replacements are found in transmembrane domains, yet these domains account for 45.0% of the 1515 amino acids that were compared. On the other hand, 45.3% of the amino acid substitutions are in extracellular regions, which comprise 29.0% of the opsin proteins. The cytoplasmic domains constitute 26.0% of the opsin proteins, and 26.8% of the amino acid substitutions occur in these domains. A χ^2 test of

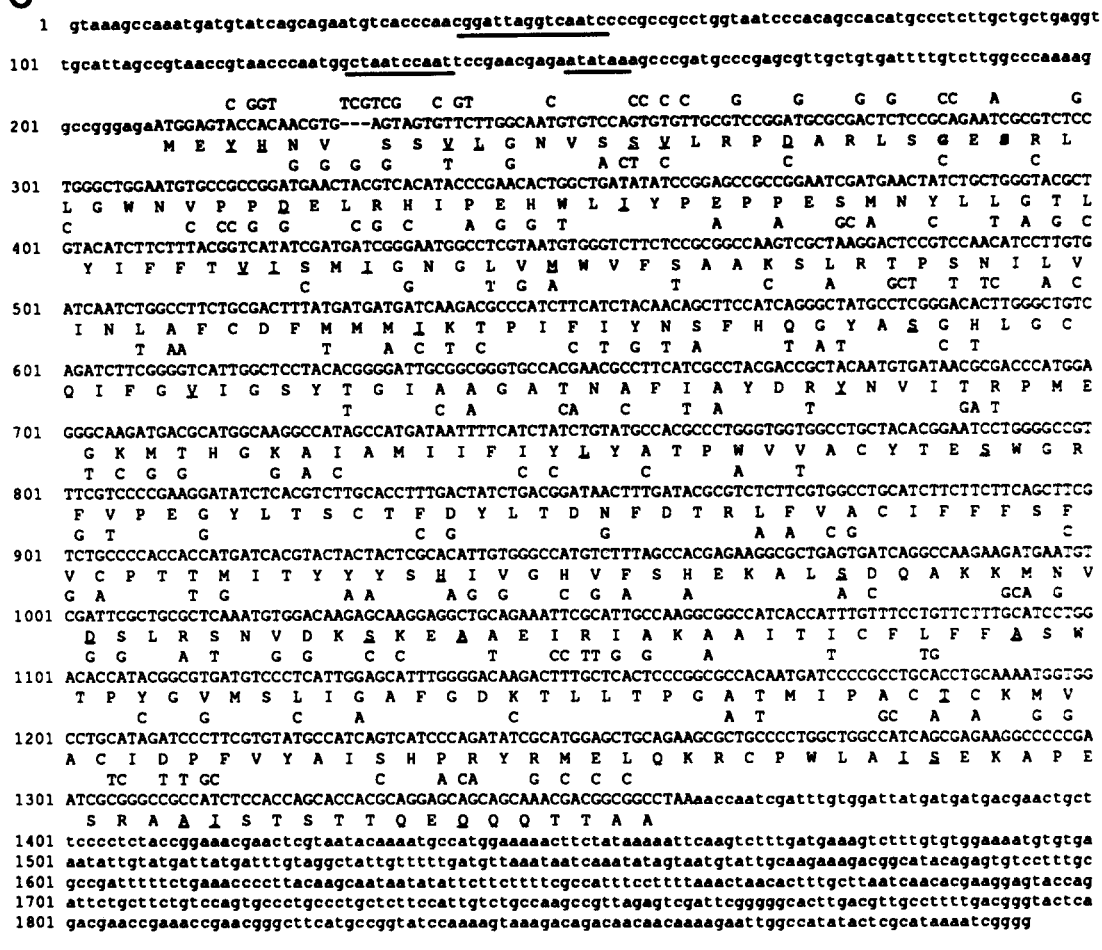

FIGURE 2.—Continued

the uniformity of amino acid replacement in the different domains shows that the distribution of amino acid replacements is clearly nonrandom ($\chi^2 = 13.38$, $P < 0.01$).

The major opsin, Rh1, has only 16 amino acid replacements between the two species. Most of the amino acid replacements at this locus are of a conservative nature, and half of the replacements are in the 49 amino acid extracellular tail. Several mutant *ninaE* alleles in *D. melanogaster* are associated with single amino acid changes in the Rh1 opsin, but none of the replacements are at sites that correspond to those *ninaE* mutations (WASHBURN and O'TOUSA 1989). Ten of the 16 replacements are at the same residues that differ between *D. melanogaster* and *Calliphora erythrocephala*, another Dipteran (HUBER *et al.* 1990). Interestingly, there are few replacements, and no charge changes, in the transmembrane domains.

The ocellar opsin, Rh2, has 36 amino acid replacements. A large number of the amino acid replacements are in the 56 amino acid extracellular tail and the 42 amino acid cytoplasmic tail. There are a total of eleven amino acid replacements in the seven putative transmembrane domains. An aspartic acid to histidine replacement at position 154 in the third transmembrane domain is particularly noteworthy because it results in a net transmembrane charge change of +2. Under one model for the function of visual pigments, the total transmembrane charge is the primary factor that determines spectral sensitivity (NATHANS 1987). The spectral sensitivity of *D. pseudoobscura* (determined for a different strain than the one that we used) ocellar photoreceptors, as well as the ocellar photoreceptors of several other Drosophila species, is very similar to that of the *D. melanogaster* receptor (CHEN *et al.* 1991).

Rh3, one of the ultraviolet rhodopsins in R7 cells (ZUKER *et al.* 1987), shows a pattern of amino acid replacement that is similar to that of Rh2. Eight of 31 amino acid replacements are in the transmembrane domains. A glutamine to histidine replacement at position 243 results in a net transmembrane charge change of +1. The 17 amino acid substitutions in Rh4, the other R7 ultraviolet receptor, are of a highly conservative nature. There are five substitutions in membrane spanning domains of Rh4, none of which results in a charge change.

DISCUSSION

Gene and genome organization of opsin genes: The major structural features that are necessary for

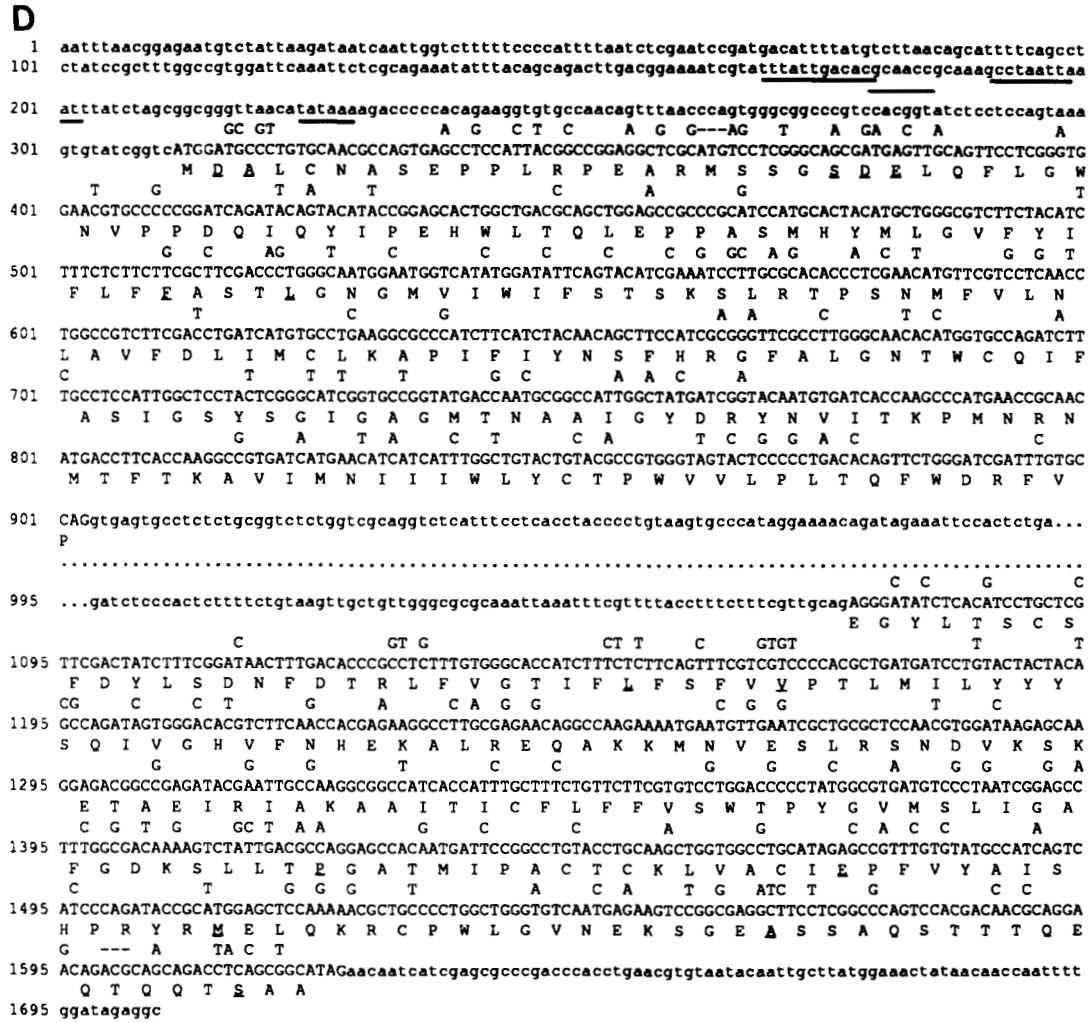

FIGURE 2.—Continued

transcription and message processing of the opsin genes are conserved between *D. pseudoobscura* and *D. melanogaster*. The positions of the introns in *Rh1*, *Rh2* and *Rh4* are precisely conserved. When paralogous comparisons are considered, only the position of intron 4 of *Rh1* and intron 2 of *Rh2* are at conserved locations (COWMAN, ZUKER and RUBIN 1986). The position of the large intron in *Rh4* is not shared with the positions of any of the other introns in Drosophila opsins. *Seven in absentia*, a gene that is nested in the intron of *D. melanogaster Rh4*, is closely linked to the 5′ exon of *Rh4* in *D. pseudoobscura*, and may also lie within the intron in this species. Detailed molecular mapping of the region will determine whether *sina* is located in the intron of *Rh4*, or if it is outside of the intron and closely linked to *Rh4*.

The opsin genes of *D. pseudoobscura* are in the same chromosome elements as their *D. melanogaster* homologs. This result is consistent with previous observations that have demonstrated that the chromosome arms of *D. pseudoobscura* and *D. melanogaster* are largely conserved with respect to genetic content (LAKOVAARA and SAURA 1982; STEINEMANN, PINSKER and SPERLICH 1984). However, the relative positions of the *Rh1*, *Rh2* and *Rh3* genes on polytene chromosome *3R* of *D. melanogaster* and *2* of *D. pseudoobscura* indicate that extensive rearrangement of this chromosome arm has occurred since these species diverged. *Rh1* and *Rh3* are in the same numbered section (92) in *D. melanogaster*, with *Rh2* in an adjacent numbered section (COWMAN, ZUKER and RUBIN 1986; FRYXELL and MEYEROWITZ 1987; MONTELL *et al.* 1987; O'TOUSA *et al.* 1985; ZUKER, COWMAN and RUBIN 1985; ZUKER *et al.* 1987). In *D. pseudoobscura*, *Rh2* and *Rh3* are in adjacent numbered sections, but *Rh1*, while on the same chromosome arm, is well separated from the other two opsins. The chromosomal location of *Rh4*, close to the telomere on the *D. pseudoobscura XR* and in the middle of *D. melanogaster 3L* also shows that there have been rearrangements of this chromsomal element. Similar rearrangements have been observed for some of the other chromosome arms (HENIKOFF and EGHTEDARZADEH 1987; SCHAEFFER and AQUADRO 1988).

Rates of amino acid replacement vary within and between opsins: Rates of amino acid replacement vary

TABLE 1

Amino acid and nucleotide evolution of Drosophila opsins

Gene	Amino acid comparisons		Nucleotide substitutions		
				Synonymous	
	N[a]	Identity (%)	Nonsynonymous Dif[b] (%)	Dif[b] (%)	Hits[c] (95% limit)
Rh1	374	95.7	3.3	26.1	2: 0.31 (0.19–0.48) 3: 0.34 (0.11–0.78) 4: 0.49 (0.35–0.67) 6: 0.73 (0.38–1.51)
Rh2	381	90.3	9.3	39.2	2: 0.78 (0.47–1.92) 3: saturated[d] 4: 0.81 (0.59–1.13) 6: 0.86 (0.48–1.68)
Rh3	383	91.1	8.6	36.0	2: 0.49 (0.31–0.77) 3: 1.70 (0.61–99.99) 4: 0.66 (0.48–0.91) 6: 1.34 (0.73–99.99)
Rh4	380	96.0	4.0	36.7	2: 0.59 (0.37–0.97) 3: 1.35 (0.53–99.99) 4: 0.68 (0.49–0.94) 6: 1.00 (0.59–1.97)

[a] N = number of amino acids.
[b] Percent of nucleotide sites that differ between the homologous opsins.
[c] Number of evolutionary substitutions per synonymous codon, determined by the method of LEWONTIN (1988). The numbers 2, 3, 4, and 6 refer to 2-, 3-, 4- and 6-fold degenerate codons.
[d] Analysis of *Rh2* sequences using two different codon usage tables (see text) resulted in infinite values for substitutions in this codon class.

significantly among the different structural and functional domains of the Drosophila ospins. There are few amino acid replacements in the transmembrane domains, and a disproportionately large number of replacements in the extracellular domains. It is not surprising that membrane spanning regions of the proteins have the lowest level of amino acid substitution, because the structural integrity of the opsins depends on the hydrophobic nature of these domains.

Despite their similar functions in phototransduction, the Drosophila opsin proteins evolve at significantly different rates. The Rh1 and Rh4 opsins are among the most conserved proteins in Drosophila, with only the heat shock protein hsp82 having a reported amino acid identity that is greater (BLACKMAN and MESELSON 1986). The Rh2 and Rh3 opsin proteins, on the other hand, fall within the middle of the range of amino acid identities reported for genes compared between these two species (RILEY 1988). The rate of amino acid replacement is not associated with the rate of synonymous nucleotide substitution. The two ultraviolet receptors (*Rh3* and *Rh4*) for example, have similar rates of synonymous nucleotide substitution, but significantly different rates of amino acid evolution (Table 1). It is also interesting to note that one of the opsins with the fewest amino acid replacements is the major opsin (Rh1).

The Rh2 ospin has the highest rate of amino acid replacement of the four opsin proteins. Eleven replacements are in transmembrane domains, including an aspartic acid to histidine substitution in the third domain that results in a net transmembrane charge difference of +2 between the Rh2 ospins of *D. melanogaster* and *D. pseudoobscura*. The fact that these two opsins have very similar spectral properties (CHEN *et al.* 1991) does not easily fit a model in which the total transmembrane charge is the primary determinant of spectral sensitivity, and is consistent with direct experimental evidence from bovine rhodopsin in which only one of several site-directed charge changes has a dramatic affect on the spectral properties of the rhodopsin (NATHANS 1990a,b; SAKMAR, FRANKE and KHORANA 1989, 1991; ZHUKOVSKY and OPRIAN 1989).

Rate of nucleotide substitution varies significantly among the opsin genes: Nucleotide substitution rates vary significantly among loci that have been compared between *D. pseudoobscura* and *D. melanogaster*. For example, RILEY (1988) found that percent synonymous site substitutions ranged from 24.2% for *Hsp*82 to 42.7% for *Gart*. Surprisingly, the variation in evolutionary rate among the four members of a single gene family, the opsins, spans a range similar to that observed among numerous genes of disparate function. The amount of synonymous nucleotide substitution in the opsin genes varies from 26.1% for *Rh1* to 39.2% for *Rh2*. Synonymous substitutions have no

TABLE 2

Codon bias and base composition of Drosophila opsin genes

Gene	*D. melanogaster*					*D. pseudoobscura*				
			Percent G+C					Percent G+C		
	F_{op}[a]	χ^2/L[b]	Entire region	Non-coding	3rd base	F_{op}[a]	χ^2/L[b]	Entire region	Non-coding	3rd base
Rh1	0.674	0.747	44.8	35.4	80.4	0.571	0.408	44.6	37.4	70.7
Rh2	0.501	0.299	44.2	33.1	68.0	0.649	0.590	53.3	45.7	82.2
Rh3	0.568	0.518	48.2	40.1	73.1	0.560	0.442	50.4	43.0	74.3
Rh4	0.576	0.411	51.2	41.6	76.0	0.603	0.428	50.6	42.1	75.9

[a] The frequency of optimal codons for both species was calculated using the optimal codons determined by D. M. SHARP (personal communication) for *D. melanogaster*.
[b] χ^2/L is the chi-square for deviation from equal use of all codons within a synonymous group. divided by the number of codons (not including methinine and tryptophan) in the gene (SHIELDS *et al.* 1988).

TABLE 3

Correlation between codon bias and base composition of the Drosophila opsin genes, and the corresponding synonymous substitution rates observed in the interspecific comparisons

	Pearson's (r) correlation coefficient	
	D. melanogaster	*D. pseudoobscura*
F_{op} *vs.* SSD[a]	−0.963*	0.609
χ^2/L *vs.* SSD	−0.965*	0.608
%GC entire region *vs.* SSD	0.291	0.991**
%GC noncoding *vs.* SSD	0.144	0.966*
%GC third base *vs.* SSD	−0.877	0.853
%GC noncoding *vs.* %GC third base	0.302	0.919

[a] SSD = synonymous site difference between homologous opsin loci.

* $P < 0.05$, ** $P < 0.01$.

effect on the amino acid sequence of the protein, so the variation in evolutionary rate must be associated with forces other than selection on protein function. Two factors that have been found to affect the rates and patterns of nucleotide substitutions in eukaryotic genomes are local base composition and codon usage (IKEMURA 1985; MARTIN and MEYEROWITZ 1988; SHARP 1989; SHARP and LI 1987; SHIELDS *et al.* 1988; WOLFE, SHARP and LI 1989). We examined the relationship between these factors and the rate of nucleotide substitution in the opsin genes.

There are two separate features of the base composition data that must be considered in this context: the relationship between base compositon in noncoding DNA and base composition at synonymous nucleotide sites; and the relationship between base composition and the rate of synonymous site evolution (SHIELDS *et al.* 1988; MORIYAMA 1992). There is significant variation in base composition among the Drosophila opsin genes, but there is only a very weak correlation between the base composition of the noncoding DNA (5′ and 3′ flanking regions and introns) and the base composition at third positions in codons. In a related study, SHIELDS *et al.* (1988) did not find a significant correlation between base composition of noncoding DNA and base composition of silent sites among 84 genes from *D. melanogaster*. If there is any mutational bias in the direction of a specific, localized, base composition in the Drosophila genome, it is not reflected in patterns of synonymous site evolution in the opsin loci.

The relationship between base composition and rate of evolution of the opsin genes is complex. There is a significant correlation between noncoding G + C content and synonomous substitution rate of the *D. pseudoobscura* opsin loci in the opsin genes, but there is no similar correlation for the *D. melanogaster* opsins. Large differences in base composition of homologous opsin loci in the two species confound attempts to identify relationships between local base composition and rate of nucleotide substitution. For example the base composition of the *Rh2* region differs significantly between *D. pseudoobscura* and *D. melanogaster*, in both noncoding DNA and in the third bases of codons. Analysis of the relationship between G + C

NH_2	out	mem	in	mem	out	mem	in	mem	out	mem	in	mem	out	mem	in
Rh1	**49**	**25**	**12**	**23**	**18**	**26**	**7**	**21**	**34**	**28**	**33**	**24**	**8**	**24**	**42**
	E2-D V6-A	M60-I						W175-C	A186-V		E269-D	T285-S			D361-E
	A7-V A9-T							+C182							A368-T
	H14-Q P17-A														
	L18-P N40-D														
Rh2	**56**	**25**	**12**	**23**	**18**	**26**	**7**	**21**	**34**	**28**	**33**	**24**	**8**	**24**	**42**
	H5-L T9-P	A66-V		S102-L		G135-A		L178-A	A197-S	L236-M	E257-D	V301-I			K342-N
	F11-L D12-A	M68-L				F153-L		V189-I	M215-Q	A247-T	G279-N				Y343-D
	L13-M H16-L					***D154-H***									I345-L
	S17-L Q21-G														F356-C
	S24-T S25-G														A368-P
	G28-R H40-P														D378-E
	W45-H M55-T														A381-D
Rh3	**53**	**25**	**12**	**23**	**17**	**26**	**12**	**21**	**29**	**28**	**34**	**24**	**9**	**24**	**42**
	S3-Y G4-H	L70-V		V109-I	L125-S	F156-Y		M181-L	T194-S	***Q243-H***	R256-S	C295-A		A326-T	L357-I
	-S7 L10-V	I80-M									E265-D				N358-S
	F11-L T16-S										N274-S				S366-R
	A17-V E21-D										T276-A				V368-A
	A26-G T28-S														A369-I
	E38-D T49-I														P377-Q
Rh4	**53**	**25**	**12**	**23**	**17**	**26**	**12**	**21**	**29**	**28**	**34**	**24**	**9**	**24**	**41**
	E2-D	C66-F								C228-V			Q307-P	D324-E	L337-M
	P3-A	V70-L								S245-N					I355-A
	+M17														Q367-T
	N21-S														T369-Q
	G22-D														+S371
	D23-G														

FIGURE 3.—Summary of amino acid replacements in each *D. pseudoobscura* opsin relative to its *D. melanogaster* homolog. The opsins have seven membrane spanning regions, with the amino end outside of the rhabdomere membrane of the photoreceptor cell and the carboxyl end in the cytoplasm. The figure is organized to highlight the structural domains of the opsin proteins: mem = transmembrane domain; out = extracellular domain; in = cytoplasmic domain. The number of amino acid residues in each domain of the proteins is in bold type. Two substitutions that result in transmembrane charge changes, one in TM III of Rh2 and one in TM V of Rh3, are shown in bold italics.

content and rate of evolution in more closely related species, where homologous loci have not diverged significantly in base composition, may provide a superior test of the hypothesis that these two factors are related.

The relationship between codon usage bias and rate of evolution of the opsin loci is also rather inconsistent. There is a significant negative correlation between codon bias of the *D. melanogaster* opsins and the synonymous site difference between homologous opsin genes. However, the correlation between codon usage bias and synonymous substitution rate is true only for the *D. melanogaster* codon bias measures. Codon usage bias for the *Rh1* and *Rh2* loci differs significantly between *D. pseudoobscura* and *D. melanogaster*, and there is no correlation between the *D. pseudoobscura* codon bias and synonymous substitution. Similar investigations (SHARP and LI 1989) have found significant correlations between codon bias and the rate of synonymous site evolution in Drosophila, so the opsin loci may represent an exception to this pattern.

Analysis of the evolution of the Drosophila opsin loci reveals several unexpected properties. The genes encode proteins which perform similar functions, but interspecific comparison of the genes reflects very different evolutionary histories. The mutational and selective forces that are responsible for the observed patterns are difficult to infer, but a few clues have been left in the DNA sequences. One of the more interesting features of the data is the decoupling of nucleotide and amino acid evolution in *Rh3* and *Rh4*, two genes that are expressed in the same cell type in the compound eye. These loci have similar base compositions, codon biases, and synonymous substitution rates in *D. melanogaster* and *D. pseudoobscura*, yet there are significantly fewer amino acid replacements in the Rh4 protein than in the Rh3 protein. These data suggest that the underlying mutational properties of these loci are similar, but that amino acid replacements have been selected against more strongly in *Rh4* than in *Rh3*.

While *Rh3* and *Rh4* appear to manifest the effects of differing selective pressures on two similar genes, analysis of the evolution of the *Rh2* locus reveals what appear to be the results of distinct mutational biases. This gene is associated with a dramatic difference in base composition, a large difference in codon usage bias, and a high rate of synonymous site substitution. The pronounced disparity in all of these properties of *Rh2* in *D. pseudoobscura* and *D. melanogaster* suggests that genomic context can influence the evolution of some loci, perhaps because of the existence of different mutational properties in different genomes or in different locations in the same genome.

Rh1 is the most slowly evolving opsin gene in Drosophila, at both the nucleotide level and the amino acid level. The slower rate of evolution of this gene, relative to the other genes in the same family, may be related to its role as the major photoreceptor molecule in the fly's visual system. The different rates of evolution among all of the opsin genes may be related to a combination of genomic factors, such as local base composition and codon usage bias, and they may result from the cumulative effects of different selection pressures. Further investigation into the evolution of these ubiquitous genes, such as the study of intraspecific nucleotide polymorphism and the inference of opsin gene phylogenies in different insects, will elucidate the factors that influence the evolution of gene families as well as the different mechanisms by which animals adapt to their visual environment.

We thank JEFF LAWRENCE for discussion and for sharing unpublished computer programs; HOWARD OCHMAN for advice and critical reading of the manuscript; PAUL SHARP for suggestions regarding analysis of codon usage bias, and for sharing data on optimal codons; and WILLIAM STARK for advice and critical reading of the manuscript. RICHARD CARTHEW and JOSEPH O'TOUSA provided clones, and STEPHEN SCHAEFFER provided the *D. pseudoobscura* genomic library. RICHARD LEWONTIN and CHUNG-I WU shared computer programs that were used for the analysis of nucleotide substitutions. DAN KRANE and two anonymous reviewers also provided essential criticism of the manuscript. This research was supported by National Institutes of Health grants GM33741 and HG00357 to D.L.H. J.P.C. was supported by National Eye Institute predoctoral Training Grant EY07108. The sequences presented in this paper were submitted to EMBL Sequence Library and have been assigned accession numbers X65877, X65878, X65879, X65880 and X65881.

LITERATURE CITED

AJIOKA, J. W., D. A. SMOLLER, R. W. JONES, J. P. CARULLI, A. E. VELLEK, D. GARZA, A. J. LINK, I. W. DUNCAN and D. L. HARTL, 1991 *Drosophila* genome project: one-hit coverage in yeast artificial chromosomes. Chromosoma **100:** 495–509.

ASHBURNER, M., 1989 *Drosophila: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y..

BEVERLY, S. M., and A. C. WILSON, 1984 Molecular evolution in *Drosophila* and the higher Diptera. II. A time scale for fly evolution. J. Mol. Evol. **21:** 1–13.

BLACKMAN, R. K., and M. MESELSON, 1986 Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila Hsp82* gene. J. Mol. Biol. **188:** 499–515.

CARTHEW, R. W., and G. M. RUBIN, 1990 *Seven in absentia*, a gene required for specification of R7 cell fate in the *Drosophila* eye. Cell **63:** 561–577.

CHEN, D.-M., W. S. STARK, J. P. CARULLI and G. BROWN, 1991 Mechanisms of ultraviolet visual sensitivity in defferent species of *Drosophila*. Invest. Opthamol. Vis. Sci. Suppl. **32:** 1152 (abstract).

COWMAN, A. F., C. S. ZUKER and G. M. RUBIN, 1986 A rhodopsin expressed in only one photoreceptor cell type of the *Drosophila* eye. Cell **44:** 705–710.

CRAXTON, M., 1991 Linear amplification sequencing, a powerful new method for sequencing DNA. Methods: A Companion to Methods in Enzymology **3:** 20–26.

DEVEREUX, J., P. HAEBERLI and O. SMITHIES, 1984 A comprehensive set of sequence programs for the VAX. Nucleic Acids Res. **12:** 387–395.

FEINBERG, A. P., and B. VOGELSTEIN, 1983 A technique for

radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

FORTINI, M. E., and G. M. RUBIN, 1990 Analysis of *cis*-acting requirements of the *Rh3* and *Rh4* genes reveals a bipartite organization to rhodopsin promoters in *Drosophila melanogaster*. Genes Dev. **4:** 444–463.

FRYXELL, K. J., and E. M. MEYEROWITZ, 1987 An opsin gene that is expressed only in the R7 photoreceptor cell of *Drosophila*. EMBO J. **6:** 443–451.

FRYXELL, K. J., and E. M. MEYEROWITZ, 1991 The evolution of rhodopsins and neurotransmitter receptors. J. Mol. Evol. **33:** 367–378.

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. **256:** 415–439.

HELMS, C., M. Y. GRAHAM, J. E. DUTCHIK and M. V. OLSON, 1985 A new method for purifying lambda DNA from phage lysates. DNA **4:** 39–49.

HENIKOFF, S., and M. K. EGHTEDARZADEH, 1987 Conserved arrangement of nested genes at the *Drosophila Gart* locus. Genetics **117:** 711–725.

HU, K. G., H. REICHERT and W. S. STARK, 1978 Electrophysiological characterization of *Drosophila* ocelli. J. Comp. Physiol. **126:** 15–24.

HUBER, A., D. P. SMITH, C. S. ZUKER and R. PAULSEN, 1990 Opsin of *Calliphora* peripheral photoreceptors R1–6: homology wih *Drosophila Rh1* and post-translational processing. J. Biol. Chem. **265:** 17906–17910.

IKEMURA, T., 1985 Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. **2:** 13–34.

LAKOVAARA, S., and A. SAURA, 1982 Evolution and Speciation in the *Drosophila obscura* group, pp. 1–59 in *The Genetics and Biology of Drosophila*, Vol. 3b, edited by M. ASHBURNER, H. L. CARSON and J. N. THOMPSON, JR. Academic Press, London.

LEWONTIN, 1988 Inferring the number of evolutionary events from DNA coding sequence differences. Mol. Biol. Evol. **6:** 15–32

LI, W.-H., and D. GRAUR, 1991 *Fundamentals of Molecular Evolution*. Sinauer Associates, Sunderland, Mass.

MARTIN, C. H., and E. M. MEYEROWITZ, 1988 Mosaic evolution in the *Drosophila* genome. Bioessays **9:** 65–69.

MARTIN, R. L., C. WOOD, W. BAEHR and M. L. APPLEBURY, 1986 Visual pigment homologies revealed by DNA hybridization. Science **232:** 1266–1269.

MENZEL, R., 1979 Spectral sensitivity and color vision in invertebrates, pp. 503–580 in *Handbook of Sensory Physiology*, Vol. VII, edited by H. AUTRUM. Springer-Verlag, Berlin.

MONTELL, C., K. JONES, C. S. ZUKER and G. M. RUBIN, 1987 A second opsin gene expressed in the ultraviolet sensitive R7 photoreceptor cells of *Drosophila melanogaster*. J. Neurosci. **7:** 1558–1566.

MORIYAMA, E. N., and T. GOJOBORI, 1992 Rates of synonymous substitution and base composition of nuclear genes of Drosophila. Genetics **130:** 855–864.

NATHANS, J., 1987 Molecular biology of visual pigments. Annu. Rev. Neurosci. **10:** 163–194.

NATHANS, J., 1990a Determinants of visual pigment absorbance: Identification of the retinylidine Schiff's base counterion in bovine Rhodopsin. Biochemistry **29:** 9746–9752.

NATHANS, J., 1990b Determinants of visual pigment absorbance: role of charged amino acids in the putative transmembrane segments. Biochemistry **29:** 937–942.

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., T. P. PIANTIDA, R. L. EDDY, T. B. SHOWS and D. S. HOGNESS, 1986 Molecular genetics of inherited variation in human color vision. Science **232:** 203–210.

NEUFELD, T. P., R. W. CARTHEW and G. M. RUBIN, 1991 Evolution of gene position: chromosomal arrangement and sequence comparison of the *Drosophila melanogaster* and *Drosophila virilis sina* and *Rh4* genes. Proc. Natl. Acad. Sci. USA **88:** 10203–10207.

O'TOUSA, J. E., E. 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

PAK, W., and S. R. GRABOWSKI, 1978 Physiology of the visual and flight systems, pp. 553–604 in *The Genetics and Biology of Drosophila*, Vol. 3a, edited by M. ASHBURNER and T. R. F. WRIGHT. Academic Press, New York.

POLLOCK, J. A., and S. BENZER, 1988 Transcript localization of four opsin genes in the three visual organs of *Drosophila*. Nature **333:** 779–782.

RILEY, M., 1988 Nucleotide sequence of the *Xdh* region in *Drosophila pseudoobscura* and an analysis of the evolution of synonymous codons. Mol. Biol. Evol. **6:** 33–52.

SAKMAR, T. P., R. R. FRANKE and H. G. KHORANA, 1989 Glutamic acid-113 serves as the retinylidine Schiff base counterion in bovine rhodopsin. Proc. Natl. Acad. Sci. USA **86:** 8309–8313.

SAKMAR, T. P., R. R. FRANKE and H. G. KHORANA, 1991 The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa. Proc. Natl. Acad. Sci. USA **88:** 3079–3083.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y..

SCHAEFFER, S., and C. F. AQUADRO, 1987 Nucleotide sequence of the alcohol dehydrogenase region in *Drosophila pseudoobscura*: evolutionary changes and evidence for ancient gene duplication. Genetics. **117:** 61–73.

SHARP, P. M., and W.-H. LI, 1987 The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. Mol. Biol. Evol. **4:** 222–230.

SHARP, P. M., and W-H. LI, 1989 On the rate of DNA sequence evolution in *Drosophila*. J. Mol. Evol. **28:** 398–402.

SHIELDS, D. C., P. M. SHARP, D. G. HIGGINS and F. WRIGHT, 1988 "Silent" sites in *Drosophila* are not neutral: evidence for selection among synonymous codons. Mol. Biol. Evol. **5:** 704–716.

STEINEMANN, M., W. PINSKER and D. SPERLICH, 1984 Chromosome homologies in the *Drosophila obscura* group of species probed by *in situ* hybridization. Chromosoma **91:** 46–53.

WASHBURN, T., and J. E. O'TOUSA, 1989 Molecular defects in *Drosophila* rhodopsin mutants. J. Biol. Chem. **264:** 15464–15466.

WOLFE, K. H., P. M. SHARP and W.-H. LI, 1989 Mutation rates differ among regions of the mammalian genome. Nature **337:** 283–285.

YOKOYAMA, S., and R. YOKOYAMA, 1989 Molecular evolution of the human visual pigment genes. Mol. Biol. Evol. **6:** 186–197.

ZHUKOVSKY, E. A., and D. D. OPRIAN, 1989 Effect of carboxylic acid side chains on the absorption maximum of visual pigments. Science **246:** 928–930.

ZUKER, C. S., A. F. COWMAN and G. M. RUBIN, 1985 Isolation and structure of a rhodopsin gene from Drosophila. 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:** 1550–1557.

Communicating editor: R. E. DENELL