ELSEVIER

Gene 230 (1999) 101–109

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Intron splice sites of *Papilio glaucus PglRh3* corroborate insect opsin phylogeny ☆

Adriana D. Briscoe *

Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA 02138, USA

Received 2 July 1998; accepted 22 December 1998; Received by W.-H. Li

Abstract

Full-length cDNA clones encoding the PglRh3 opsin from the tiger swallowtail butterfly *Papilio glaucus* were isolated from cDNA synthesized from adult head tissue total RNA. This cDNA consists of 1679 nucleotides and contains a single open reading frame predicted to be 379 amino acids in length. PCR amplification of genomic DNA with primers spanning the coding region yielded a single 2760 bp fragment which was sequenced. The *PglRh3* gene has nine exons and eight introns, four of which are in unique locations relative to the positions of introns in other known insect opsin sequences. Phylogenetic analyses of amino acid and nucleotide sequence data places PglRh3 within a clade of insect visual pigments thought to be sensitive to long wavelengths of light. The genomic structure of *PglRh3* is the first characterized from a member of this opsin clade. Three *PglRh3* intron positions are shared with *Drosophila Rh1*, and one of these is also shared with *Drosophila Rh2*. By contrast, none of the known intron locations in a clade of anciently diverged ultraviolet- and blue-sensitive visual pigments are shared by *P. glaucus PglRh3, Drosophila Rh1* or *Rh2*. The placement of introns within opsin genes therefore independently supports the clustering of a putatively long-wavelength-sensitive clade with a clade of blue–green-sensitive visual pigments.

Keywords: Butterfly; Insect opsin phylogeny; Intron splice sites; *Papilio glaucus*; *PglRh3*; Visual pigments

1. Introduction

Comparative studies of the evolutionary relationships among insect visual pigments have identified three to four major opsin lineages (e.g. Briscoe, 1998; Townson et al., 1998). Within each of these lineages, the spectral properties of the visual pigments appear to evolve somewhat conservatively. That is to say, ultraviolet (UV)-sensitive visual pigments that have been physiologically characterized by transgenic expression fall into the same clade as other similarly characterized UV-sensitive pigments (Feiler et al., 1992; Townson et al., 1998). The same pattern holds for the blue-, blue–green- and green- or long-wavelength-sensitive pigments that have been expressed transgenically so far (Table 1) (Zuker et al., 1988; Feiler et al., 1988; Townson et al., 1998; Britt, personal communication).

Phylogenies based upon opsin amino-acid (aa) sequences group the UV- with the blue-sensitive opsin clade, and the blue–green- with the long-wavelength-sensitive clade. The clustering of the UV- and blue-sensitive clades is supported by a shared intron splice site. By contrast, intron splice sites are found in completely different locations in the blue–green-sensitive clade (Carulli et al., 1994; Bellingham et al., 1997), and are unknown in the long-wavelength-sensitive clade. This study supports the relationship among these clades by providing the first intron–exon structure of a member of an insect clade of opsins thought to be sensitive to long wavelengths of light.

Visual pigments are composed of an opsin protein bound to a retinal derivative chromophore by a Schiff base linkage (Bownds, 1967). Opsin proteins are encoded by members of the G protein-coupled receptor gene superfamily (Fryxell and Meyerowitz, 1991). Most organisms express only one kind of chromophore and

* Present address: Department of Molecular and Cellular Biology, 1007 E. Lowell/Life Sciences South, University of Arizona, Tucson, AZ 85721, USA. Tel.: +1 520-626-2368; Fax: +1 520-621-3709; e-mail: abriscoe@u.arizona.edu

☆ The nucleotide sequence data reported will appear in the GenBank Nucleotide Sequence Database under the accession numbers AF067080 (*PglRh3* cDNA) and AF098283 (*PglRh3* genomic DNA).

Abbreviations: aa, amino acid(s); UTR, untranslated region; UV, ultraviolet.

0378-1119/99/$ – see front matter
PII: S0378-1119(99)00027-X

Table 1
Insect opsins which have been expressed transgenically in *Drosophila*, and physiologically characterized by spectral sensitivity measurements

Species	Abbreviation	λ max	Reference
Apis mellifera	Bee UV	353 nm	Townson et al., 1998
A. mellifera	Bee blue	439 nm	Townson et al., 1998
Drosophila melanogaster	Rh1	480 nm	Zuker et al., 1988; Feiler et al., 1988
D. melanogaster	Rh2	420 nm	Zuker et al., 1988; Feiler et al., 1988
D. melanogaster	Rh3	345 nm	Feiler et al., 1992
D. melanogaster	Rh4	375 nm	Feiler et al., 1992
D. melanogaster	Rh5	[a]	(S.M. Britt, personal communication)
D. melanogaster	Rh6	[a]	(S.M. Britt, personal communication)

[a] Spectral sensitivity measurements of photoreceptor cells expressing these opsins in transgenic *Drosophila* are consistent with prevoius spectral sensitivity measurements of photoreceptor cells expressing these opsins in normal flies (see Hardie, 1985). The R8 photoreceptor cell subtype in *Drosophila* which expresses Rh5 (Chou et al., 1996; Papatsenko et al., 1997) is thought to be blue-sensitive, while the R8 photoreceptor cell subtype which expressed Rh6 (Huber et al., 1997) is thought to be green-sensitive.

changes in the use of chromophores tend to be conservative, which means that differences in the maximal spectral absorbance (λ max) between visual pigments are due primarily to changes in the aa sequences of the opsin. In general, the wavelength of light at which a visual pigment maximally absorbs conforms closely to the electrophysiologically measured peak spectral sensitivity of the photoreceptor expressing that pigment, although such peaks can be shifted due to the effects of oil droplets (e.g. Bowmaker et al., 1997), sensitizing pigments (Gärtner and Towner, 1995), or screening pigments (Gribakin, 1988). The aa sequences of the opsins identified from some butterflies, including members of the genus *Papilio*, are of particular interest since they have at least two spectral classes of long-wavelength receptor (Bernard, 1979; Arikawa et al., 1987), one of which is considerably more red-shifted (600 nm) than the long-wavelength-sensitive receptor of most insects (520–530 nm) (for a brief review see Chittka, 1996). Previous work on *Papilio glaucus* reported the cloning of six opsin-encoding cDNA fragments, PglRh1–6, four of which (PglRh1–4) fall in the long-wavelength-sensitive clade (Briscoe, 1998). Isolation of the full-length sequences of *P. glaucus PglRh3* is an important first step towards determining the aa substitutions that account for the spectral sensitivity differences between butterfly long-wavelength-sensitive visual pigments. The coding region of *Papilio PglRh3* will be used to make a construct for transgenic expression and physiological characterization in an already existing *Drosophila* transgenic system (Feiler et al., 1988, 1992; Townson et al., 1998).

2. Materials and methods

2.1. *Cloning and sequencing of the* PglRh3 *cDNA*

Partial cDNA clones of *PglRh3* were isolated as previously described (Briscoe, 1998). A gene-specific reverse primer (P1R1, 5′-TGT ATT TCG GGT GGC TGA T-3′) was designed using Oligo 4.0. A pool of cDNA was sythesized from head tissue total RNA using the Marathon cDNA Isolation kit (Clontech) and used in 5′ RACE PCR with the primer P1R1 and the adaptor primer AP1 (Clontech) (PCR conditions: 1 min at 94°C, then 30 cycles of 30 s at 94°C, 30 s at 55°C, 3 min at 68°C). To minimize misincorporations of nucleotides, and verify previously obtained overlapping sequences, full-length sequences were amplified from cDNA using gene-specific forward and reverse primers (P1F4, 5′-CCG TGG TAG AAG CAA AAG AC-3′ and P1R4, 5′-GAA TCA GGT CCA GGT AAC G-3′), and the Advantage cDNA PCR kit (Clontech) which contains Klen*Taq*-1 DNA polymerase, and a small amount of a second proofreading polymerase, Deep Vent® (PCR conditions are modified from Turbeville et al., 1997: 1 min at 94°C, then 35 cycles of 30 s at 94°C, 3 min at 62°C, 3 min at 68°C). PCR products were incubated for 10 min at 72°C with 0.5 μl *Taq* DNA polymerase (Promega) to add A-overhangs, gel purified (Geneclean II kit, Bio 101), cloned into pCR2.1-TOPO plasmid (TOPO TA cloning, Invitrogen) and sequenced (Dye Primer and Dye Terminator Cycle Sequencing Ready Reaction kits, Perkin-Elmer).

2.2. *Isolation of* PglRh3 *genomic DNA*

Genomic DNA was extracted from the same adult female butterfly from which the head tissue total RNA was isolated. Approximately 100 mg thorax tissue was homogenized in 400 μl 50 mM Tris–Cl (pH 8.0)/2% SDS and transferred to a 1.5 ml tube. Two microliters of proteinase K was added, and then incubated overnight at 37°C. 100 μl saturated NaCl was added, the tube was vortexed, and placed on ice for 35 min, and then spun at 13 000 rpm for 15 min at 4°C in a Micro Centrifugette 4214 (ALC). The genomic DNA was then ethanol precipitated. Primers spanning the coding region of *PglRh3*, P1F4 and P1R4, were used in a PCR reaction

with genomic DNA as template (PCR conditions: 1 min at 94°C, then 35 cycles of 10 s at 94°C, 1 min at 62°C and 10 min at 62°C).

2.3. Phylogenetic analysis

A Blast search was used to identify insect opsin sequences (Altschul et al., 1990) which were downloaded and aligned. Alignments were made first on the aa sequences with ClustalW (Thompson et al., 1994), manually adjusted, and then the nucleotide sequences were aligned accordingly. Phylogenetic analysis was implemented using maximum parsimony and maximum-likelihood. Because maximum-likelihood (Felsenstein, 1981) is computationally intensive for more than about 20 sequences, phylogenies were first generated on the larger data set of 29 opsin aa sequences (343 characters, corresponding to aa 26–365 in PglRh3) by maximum parsimony (Swofford, 1991). Maximum parsimony was implemented with a G protein-coupled receptor stepmatrix, which weights aa substitutions according to their occurrence in a data set of 125 receptor sequences (Rice, 1994). These preliminary gene trees were used to pare down the final data set into 17 taxa representative of only the deepest nodes. For the maximum-likelihood analysis of the nucleotide sequences of these 17 taxa (981 characters), the transition:transversion rate ratio and shape parameter of the gamma distribution were estimated from the data, with four rate classes, and empirical base frequencies using the Hasegawa–Kishino–Yano (1985) model of evolution. Parsimony and maximum-likelihood analyses were performed using test versions of PAUP*4d64 (Swofford, 1991).

3. Results and discussion

3.1. Cloning of PglRh3 cDNA

12 *PglRh3* clones were isolated. Six of these clones were from two independent 3′ RACE PCR reactions, three from a single 5′ RACE PCR, and three from a single PCR reaction using primers P1F4 and P1R4. PCR products amplified with primers P1F4 and P1R4 yielded clones spanning the coding region, overlapping with the 3′ and 5′ RACE products. Including 106 bp of 5′ untranslated leader (UTR) and 433 bp of 3′ UTR, a total of 1679 bp of *PglRh3* cDNA was cloned and sequenced in both directions (GenBank accession number AF067080). Two in-frame stop codons are present at −57 and −69 nucleotides from the putative transcription initiation start site, AUGG. One polyadenylation signal is also present at bp 1645–1650.

3.2. Characterization of PglRh3 intron and exon structure

A single 2760 bp PCR product (Fig. 1) was isolated, cloned and sequenced (GenBank accession number AF098283). The gene contains nine introns, ranging from 68 to 641 bp in length (Fig. 2A). All intron splice junctions begin with GT and end with AG. Introns 2 (located between cDNA bp 240 and 241), 4 (bp 479 and 480), 5 (bp 693 and 694) and 8 (bp 1119 and 1120) interrupt an aa. Introns 3 (bp 361 and 362), 6 (bp 841 and 842) and 7 (bp 952 and 953) fall between two aa (85 and 86; 245 and 246; 282 and 283), and intron 1 (bp 98 and 99) is located in the 5′ UTR −8 nucleotides from the start codon. Introns 5, 6 and 8 in *Papilio PglRh3* are found in *Drosophila Rh1* and, intron 8 is in *Drosophila Rh2* (Fig. 2B).

3.3. Conserved PglRh3 amino-acid motifs and residues

PglRh3 shares a number of structural features with other insect visual pigments and G protein-coupled receptors. The seven transmembrane domains identified in other insect opsins by hydropathicity analysis (Kyte and Doolittle, 1982) are retained in the PglRh3 sequence. Conserved aa residues that are important in determining rhodopsin function are present as well. For instance, Lysine^{296} in bovine rhodopsin, the residue in transmembrane domain (TM) 7 to which the chromophore is bound by a protonated Schiff base linkage (Bownds, 1967), is found in PglRh3 at aa position 326 [see Chang et al. (1995) for an alignment of vertebrate and *Drosophila* opsin sequences]. An essential disulfide bridge in bovine rhodopsin (Karnik et al., 1988), connecting TM3 to the second intracellular loop, is conserved in PglRh3 (Cys$^{129–206}$). Leucine (aa 87) and asparagine (aa 92) residues in the cytoplasmic loop connecting TM1 and 2, conserved across G protein-coupled receptors, which have been shown to be critical in the synthesis of *Drosophila* Rh1, are also conserved. Mutations of these residues result in the maturation of $<2\%$ of rhodopsin transcripts (Bentrop et al., 1997). A triplet of aa which agree with the consensus site for glycosylation (N–X–S/T) in *Drosophila* opsins (O'Tousa, 1992) can be found in PglRh3 (Asn^{26}). A highly conserved G protein coupled-receptor motif (E/D R) (Baldwin et al., 1997) in the portion of TM3 adjacent to the cytosolic loop, involved in G protein binding/activation (Kim et al., 1997) is also present. Finally, there are a number of serines and threonines (Bownds, 1967) at the C-terminus which are candidate sites of phosphorylation in photoactivated rhodopsin by rhodopsin kinase (Ohguro et al., 1996; Thurmond et al., 1997).

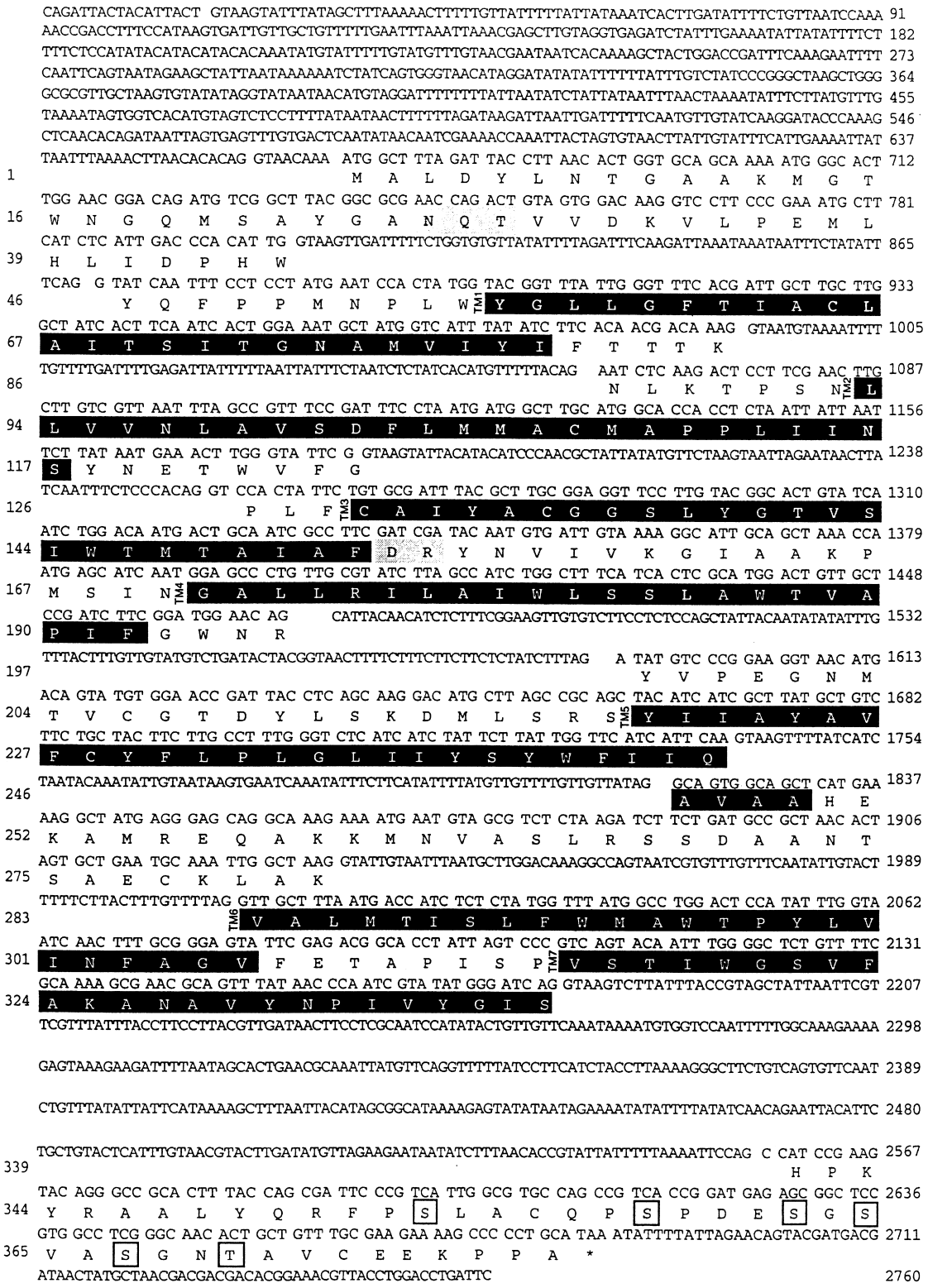
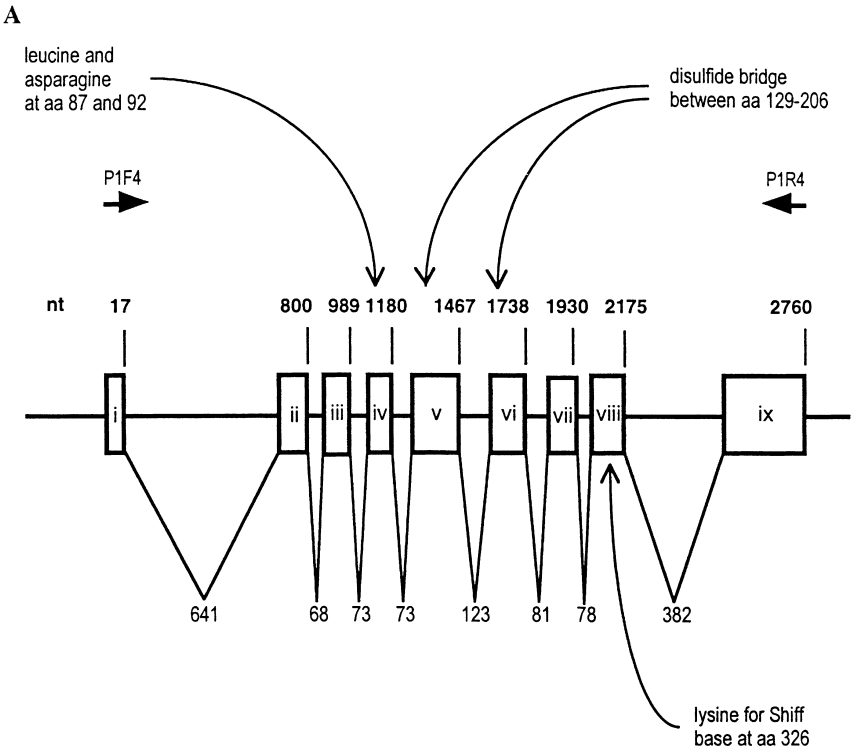


Fig. 1. Nucleotide and deduced aa sequence of *P. glaucus PglRh3* genomic DNA. The gDNA is 2760 bp in length and encodes a protein of 379 aa. Numbers on the left-hand side are for aa and those at the right are for nucleotides. The first grey box indicates a possible glycosylation site, NQT (O'Tousa, 1992), the second, E/DR, a site of G protein binding/activation (Baldwin et al., 1997). Black boxes designate seven potential transmembrane domains. Open boxes identify a number of potential phosphorylation sites (Ohguro et al., 1996; Thurmond et al., 1997). GenBank accession number: AF098283.



B

intron phase	2	0	2	1	0	2	0
aa	77	79-80	190	194	239-40	332	335-6
Dmel Rh3	FSAAKSLR	.....	WGRFVPEGYLT	.....	IVGHVFSHEK	.....	AISHPRY
Dmel Rh4	FSTSKSLR	.....	WDRFVPEGYLT	.....	IVGHVFSHEK	.....	AISHPRY
Bee Blue	FSTSKSLR	.....	WGRYTTEGFLT	.....	LLSSIRNHEK	.....	AINHPRY
Dmel Rh5	FSTSKSLR	.....	WGRYQPEGFLT	.....	LFTHVRVHEK	.....	ATSHPKY
Pgl Rh3	FTTTKNLK	.....	WNRYVPEGNMT	.....	IIQAVAAHEK	.....	GISHPKY
Dmel Rh1	FATTKSLR	.....	WSRYVPEGNLT	.....	IIAAVSAHEK	.....	GISHPKY
Dmel Rh2	FGGTKSLR	.....	WSAYVPEGNLT	.....	IIAAVAAHEK	.....	GISHPKY

Fig. 2. (A) Genomic structure of *PglRh3*. Nine exons and eight introns are present in the 2760 bp fragment amplified from genomic DNA using primers P1F4 and P1R4 (see Section 2.1 for primers). Lower-case roman numerals are used to label the exons represented by open boxes. Primer locations and orientations are indicated by black arrows. Bold numbers indicate the nuclotide site marking the 3′ boundary of exons i–ix. Bold numbers are relative to the genomic DNA sequence shown in Fig. 1. Plain letters below the boxes indicate the size of each of the eight introns in nucleotides. The putative *PglRh3* start codon is located at nucleotide 668 in exon ii. The first in-frame stop codon follows at nucleotide 2682 in exon ix. Note: 10 bp of 5′ UTR and 294 bp of 3′ UTR present in the cDNA sequence (GenBank accession number AF067080) lie outside the genomic region amplified by P1F4 and P1R4. (B) Amino acid alignment showing the location of introns shared between paralogous (related by gene duplication) opsin genes. Numbers above the alignment are relative to *Drosophila* Rh1. As shown in Fig. 4, Dmel *Rh4,* Bee blue and Dmel Rh5 share a phase-one intron located at aa 194, Dmel Rh5 and bee blue share a phase-two intron at aa 77, Pgl *Rh3* and Dmel *Rh1* share phase-two introns at aa 190 and 332, and a phase-zero intron between aa 239 and 240. Dmel *Rh2* shares the same phase-two intron with them both at aa 332.

3.4. Phylogenetic analysis of insect opsins

The maximum parsimony analysis of the larger data set (Fig. 3) is consistent with the maximum-likelihood analysis of the smaller data set (Fig. 4) at all nodes except one, where the Bee UV–Moth 2 clade collapses into a basal polytomy in the maximum parsimony reconstruction. PglRh3 is most closely related to the *Manduca sexta* Moth 1 (MANOP3 by the nomenclature in Chase et al., 1997) opsin, in both maximum parsimony and maximum-likelihood reconstructions, and falls within a clade of visual pigments thought to be sensitive to long wavelengths of light. This sequence similarity suggests that PglRh3 encodes a long-wavelength-sensitive visual pigment. The evidence for this preliminary assignment of function is indirect, however. Eight insect opsins have been transgenically expressed and physiologically characterized (Table 1): two UV-sensitive visual pigments of *D. melanogaster*; Rh3 and Rh4 (Feiler et al., 1992); the UV-sensitive honey-bee visual pigment (Bee UV)

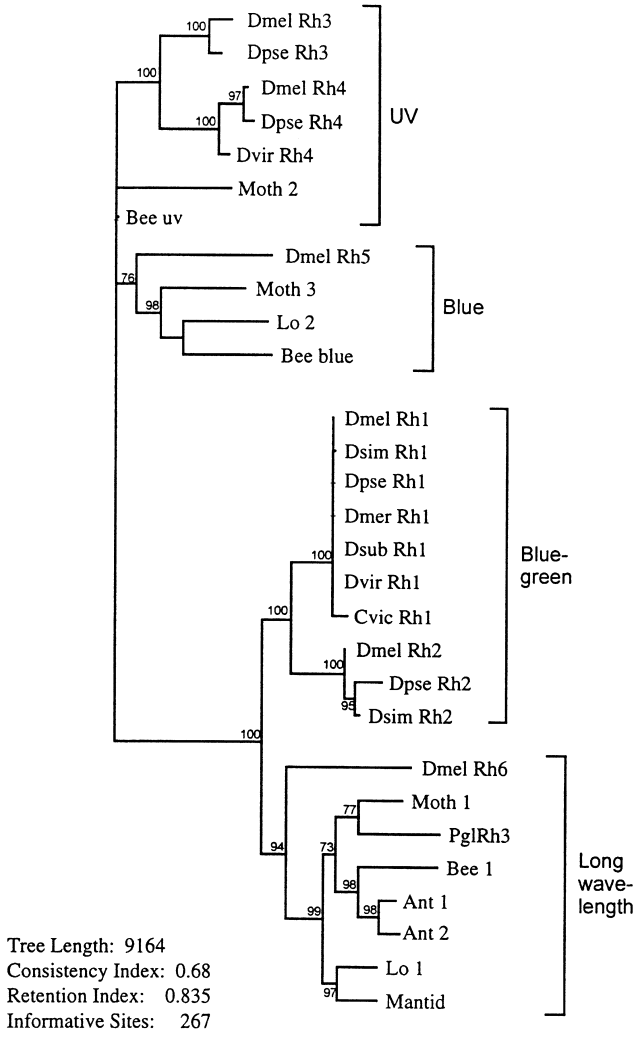


Fig. 3. Maximum parsimony phylogeny of insect opsins using a G protein-coupled receptor derived stepmatrix (Rice, 1994). An initial heuristic search resulted in a single island with five equally parsimonious trees. The tree shown is a bootstrap 50% majority-rule consensus tree. Numbers below the nodes are the number of replicates out of 1000 in which the clade was represented. Nodes with bootstrap support less than 70% are unlabelled. Tree–bisection–reconnection (TBR) branch-swapping and random sequence addition were used. Permission to publish was kindly provided by D. Swofford. Abbreviations and GenBank accession numbers: *Apis mellifera* (Bee UV, AF004169; Bee blue, AF004168; Bee 1, U26026); *Calliphora vicina* (Cvic Rh1, M58334); *Camponotus abdominalis* (Ant 1, U32502); *Cataglyphis bombycina* (Ant 2, U32501); *Drosophila melanogaster* (Dmel Rh1, K02315; Dmel Rh2, M12896; Dmel Rh3, M17718; Dmel Rh4, M17730; Dmel Rh5, U67905; Dmel Rh6, Z86118); *Drosophila pseudoobscura* (Dpse Rh1, S45126; Dpse Rh2, S45127; Dpse PglRh3, S45849; Dpse Rh4, S45850); *Drosophila subobscura* (Dsub Rh1, AF025813); *Drosophila virilis* (Dvir Rh4, M77281); *P. glaucus* (Pgl Rh3, AF067080); *Schistocerca gregaria* (Lo 1, X80071; Lo 2, X80072); *Sphodromantis* (Mantid, X71665). GSDB accession numbers: *Manduca sexta* (Moth 1, 76082; Moth 2, 109852; Moth 3, 1249561); *Drosophila simulans* (Dsim Rh1, 632784). *Drosophila mercatorum*(Dmer Rh1), *D. simulans* (Dsim Rh2), and *D. virilis* (Dvir Rh1) sequences were from Carulli et al. (1994).

(Townson et al., 1998); the honey-bee blue-sensitive visual pigment (Townson et al., 1998); the two blue–green-sensitive visual pigments of *D. melanogaster* Rh1 and Rh2 (Feiler et al., 1988; Zuker et al., 1988), and the blue- and green-sensitive visual pigments of *D. melanogaster* Rh5 and Rh6 (S. Britt, pers. comm.) (see also Carulli et al., 1994 for spectral sensitivity measurements in other species of *Drosophila*) (Table 1 and Fig. 4). An examination of the distribution of these pigments on the opsin phylogeny reveals that physiologically similar visual pigments cluster together. The UV-, blue-, and blue–green-sensitive opsins all fall together in clades outside the long-wavelength clade to which *P. glaucus* PglRh3 belongs, while the green-sensitive *D. melanogaster* Rh6 falls in the same clade. Because opsins are members of a multigene family, it is possible that PglRh3 encodes a visual pigment with an extremely different sensitivity from other members of this clade. While this possibility cannot be eliminated in the absence of physiological characterization, it is somewhat diminished by the identification of *P. glaucus* opsins which fall in the UV- (PglRh5) and blue- (PglRh6) sensitive clades (Briscoe, 1998). These molecular data are in close agreement with physiological studies (Horridge, 1983; Matic, 1983; Arikawa et al., 1987) of other members of the genus which have identified UV- and blue-sensitive photoreceptor cells in the retina. Northern blots of mRNA show positive expression of PglRh3 in the retina, suggesting a visual function (data not shown), and in combination, this evidence suggests that PglRh3 may underlie the class of either green- or red-sensitive photoreceptor cells in *Papilio*. Futher work using in situ hybridization and immunohistochemistry will clarify this role.

3.5. Genomic structure of PglRh3 and corroboration of opsin phylogeny

The genomic structures of several paralogous insect opsin genes are shown in Fig. 4. The two deepest nodes of the opsin phylogeny are supported by the sharing of a single intron by the UV- and blue-sensitive clades and by the sharing of three introns by the blue–green-sensitive and long-wavelength-sensitive opsin clades. None of the introns known from the blue–green-sensitive or long-wavelength sensitive clades are shared with members of the UV- or blue-sensitive clade. The reasons for the absence of shared introns between the deepest branches of the opsin gene family are unclear, but may be due to the loss of introns or restricted phylogenetic sampling. Five of the seven known genomic structures are of *Drosophila* loci. Recent studies of the deletion rate of DNA in "dead-on-arrival" copies of transposable elements suggest that *Drosophila* has a high rate of

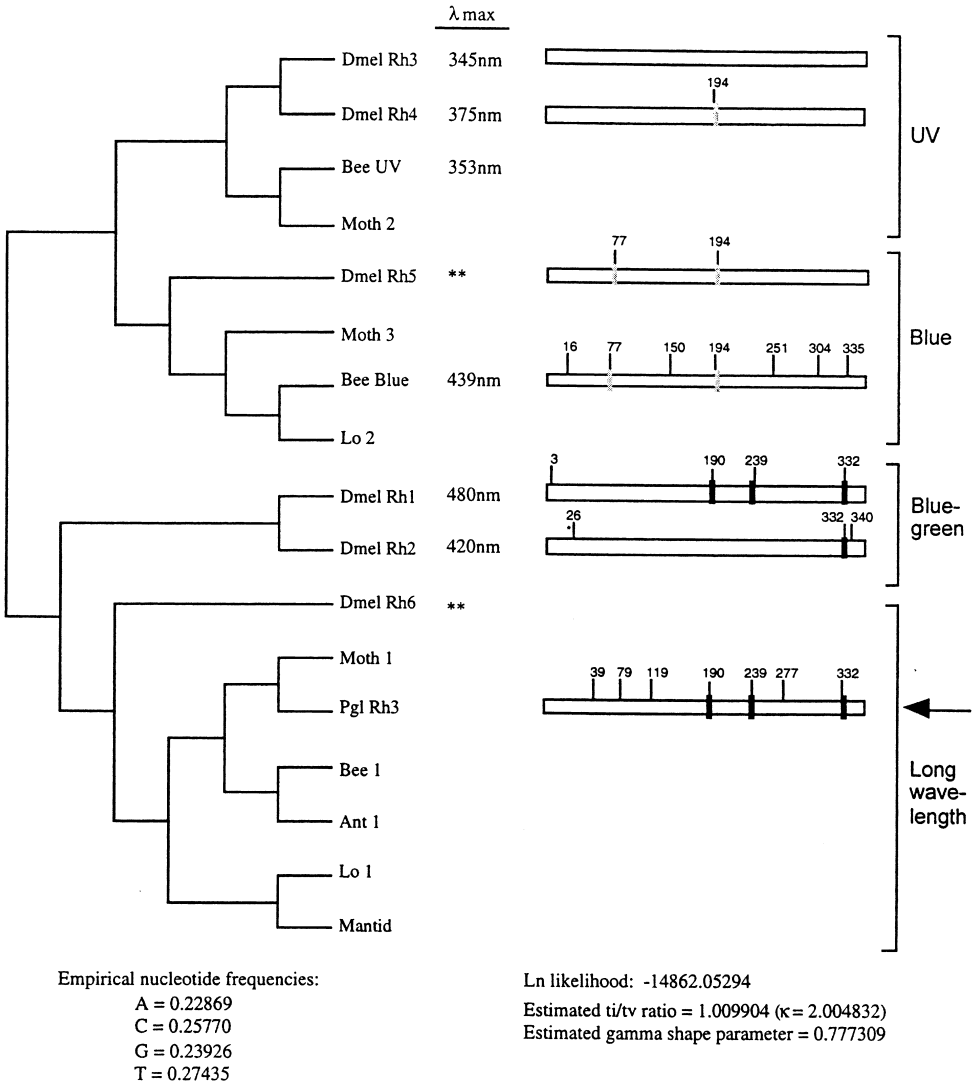


Fig. 4. Maximum-likelihood phylogeny of opsin nucleotide sequences, spectral sensitivities of transgenically expressed opsins, and genomic structures of representative taxa. Black numbers above open rectangles: location of introns relative to *Drosophila melanogaster* Rh1 aa sequence. Note: the genomic structures of *D. pseudoobscura Rh4* are identical to their *D. melanogaster* orthologues (Carulli et al., 1994). *D. melanogaster Rh5* is characterized in Chou et al. (1996). The intron–exon structure of the *Apis mellifera* blue opsin gene is given in Bellingham et al. (1997) (see also Townson et al., 1998). Black arrow: *P. glaucus PglRh3* genomic structure (intron 1 is not shown because it falls 5′ to the coding region). Grey and black bold lines: intron(s) shared by members of the UV- and blue-sensitive opsin clade (grey) and by members of the blue–green- and long-wavelength-sensitive opsin clade (black) (see Discussion, Sections 3.4 and 3.5). Asterisk: the *Rh1* genes of *D. virilis* and *D. mercatorum* share an intron at aa 26 with the *Drosophila* Rh2 clade, although the intron at this position has been lost in the *D. melanogaster* and *D. pseudoobscura* genes. Double asterix: see Table 1 for spectral information.

DNA loss (Petrov et al., 1996). Sampling of the genomic structures of other members of the UV- and blue-sensitive opsin clades, such as *P. glaucus PglRh5* and *PglRh6* (Briscoe, 1998), may reveal the presence of introns shared by members of the blue–green-sensitive and long-wavelength-sensitive opsin clades. The likelihood of such a finding is highlighted by the report of Bellingham et al. (1997) that the human blue-sensitive visual pigment shares non-overlapping introns with both the honey bee blue-sensitive opsin (aa 335) and *Drosophila Rh1* genes (aa 190).

4. Conclusions

1. *P. glaucus* cDNA (1679 bp) encoding an opsin protein, PglRh3, was isolated from adult head tissue and molecularly characterized.
2. Analysis of the genomic structure of *PglRh3* reveals the presence of nine exons and eight introns, three of which (5, 6, and 8) are shared with other known insect opsins.
3. Phylogenetic analyses indicate that *P. glaucus* PglRh3 is a member of an insect long-wavelength opsin clade,

and is most closely related to *Manduca sexta* MANOP3 (Chase et al., 1997). The insertion sites of introns 5, 6 and 8 are conserved between *P. glaucus PglRh3* and *Drosophila Rh1*. *Drosophila Rh2* also shares the insertion site of intron 8 with *P. glaucus PglRh3* and *Drosophila Rh1*. The conservation of these three intron locations independently supports maximum parsimony aa and maximum-likelihood nucleotide sequence-based phylogenies which group a clade of insect blue–green-sensitive opsins with the clade of long-wavelength-sensitive opsins.

Acknowledgements

I thank J. Staton for suggesting the long PCR protocol, A. Berry, D. Haig, E. Kellogg, T. Steen and two anonymous reviewers for their comments on this paper, R.C. Lewontin and N.E. Pierce for their support of this work, K. Horton for technical help in preparing the manuscript, C. Desplan for discussions and S. Britt for sharing unpublished data. This work was supported by a National Science Foundation Dissertation Improvement grant and a Sigma-Xi Grant-in-Aid of Research to A.D.B. and a National Institutes of Health grant (GM21179) to R.C. Lewontin. A.D.B. is a Howard Hughes Medical Institute PreDoctoral Fellow.

References

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., Eugene, W., Lipman, D. J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Arikawa, K., Inokuma, K., Eguchi, E., 1987. Pentachromatic visual system in a butterfly. Naturwissenschaften 74, 297–298.

Baldwin, J.M., Schertler, G.F.X., Unger, V.M., 1997. An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors. J. Mol. Biol. 272, 144–164.

Bellingham, J., Wilkie, S.E., Morris, A.G., Bowmaker, J.K., Hunt, D.M., 1997. Characterisation of the ultraviolet-sensitive opsin gene in the honey bee, *Apis mellifera*. Eur. J. Biochem. 243, 775–781.

Bentrop, J., Schwab, K., Pak, W.L., Paulsen, R., 1997. Site-directed mutagenesis of highly conserved amino acids in the first cytoplasmic loop of *Drosophila* Rh1 opsin blocks rhodopsin synthesis in the nascent state. EMBO J. 16, 1600–1609.

Bernard, G. D., 1979. Red-absorbing visual pigment of butterflies. Science 203, 1125–1127.

Bowmaker, J.K., Heath, L.A., Wilkie, S.E., Hunt, D. M., 1997. Visual pigments and oil droplets from six classes of photoreceptor in the retinas of birds. Vision Res. 37, 2183–2194.

Bownds, D., 1967. Site of attachment of retinal in rhodopsin. Nature 216, 1178–1181.

Briscoe, A. D., 1998. Molecular diversity of visual pigments in the butterfly *Papilio glaucus*. Naturwissenschaften 85, 33–35.

Carulli, J.P., Chen, D.-E., Stark, W.S., Hartl, D.L., 1994. Phylogeny and physiology of *Drosophila* opsins. J. Mol. Evol. 38, 250–262.

Chang, B.S., Crandall, K.A., Carulli, J.P., Hartl, D.L., 1995. Opsin phylogeny and evolution: a model for blue shifts in wavelength regulation. Mol. Phylogenet. Evol. 4, 31–43.

Chase, M., Bennett, R.R., White, R.H., 1997. Three opsin encoding cDNAs from the compound eye of *Manduca sexta*. J. Exp. Biol. 200, 2469–2478.

Chittka, L., 1996. Does bee color vision predate the evolution of flower color? Naturwissenschaften 83, 136–138.

Chou, W.H., Hall, K.J., Wilson, D.B., Wideman, C.L., Townson, S.M., Chadwell, L.V., Britt, S.G., 1996. Identification of a novel *Drosophila* opsin reveals specific patterning of the R7 and R8 photoreceptor cells. Neuron 11, 1101–1115.

Feiler, R., Bjornson, R., Kirschfeld, K., Mismer, D., Rubin, G.M., Smith, D.P., Socolich, M., Zuker, C.S., 1992. Ectopic expression of ultraviolet-rhodopsins in the blue photoreceptor cells of *Drosophila*: visual physiology and photochemistry of transgenic animals. J. Neurosci. 12, 3862–3868.

Feiler, R., Harris, W.A., Kirschfeld, K., Wehrhahn, C., Zuker, C.S., 1988. Targeted misexpression of a *Drosophila* opsin gene leads to altered visual function. Nature 333, 737–741.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Fryxell, K.J., Meyerowitz, E.M., 1991. The evolution of rhodopsins and neurotransmitter receptors. J. Mol. Evol. 33, 367–378.

Gärtner, W., Towner, P., 1995. Invertebrate visual pigments. Photochem. Photobiol. 62, 1–16.

Gribakin, F.G., 1988. Photoreceptor optics of the honeybee and its eye colour mutants: the effects of screening pigments on the long-wave subsystem of colour vision. J. Comp. Physiol. 164, 123–140.

Hardie, R.C., 1985. Functional organization of the fly retina. In: Autrum, H., Ottoson, D., Perl, E.R., Schmidt, R.F., Shimazu, H., Willis, W.D. (Eds.), Progress in Sensory Physiology. Springer, Berlin, pp. 1–79.

Hasegawa, M., Kishino, K., Yano, T., 1985. Dating the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Horridge, G.A., 1983. Single electrode studies on the retina of the butterfly *Papilio*. J. Comp. Physiol. 150, 271–294.

Huber, A., Schulz, S., Bentrop, J., Groell, C., Wolfrum, U., Paulsen, R., 1997. Molecular cloning of *Drosophila* Rh6 rhodopsin: the visual pigment of a subset of R8 photoreceptor cells. FEBS Lett. 406, 6–10.

Karnik, S.S., Sakmar, T.P., Chen, H.B., Khorana, H.G., 1988. Cysteine residues 110 and 187 are essential for the formation of correct structure in bovine rhodopsin. Proc. Natl. Acad. Sci. USA 85, 8459–8463.

Kim, J.M., Altenbach, C., Thurmond, R.L., Khorana, H.G., 1997. Structure and function in rhodopsin: rhodopsin mutants with a neutral amino acid at E134 have a partially activated conformation in the dark state. Proc. Natl. Acad. Sci. USA 94, 14273–14278.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.

Matic, T., 1983. Electrical inhibition in the retina of the butterfly *Papilio*, I: Four spectral types of photoreceptors. J. Comp. Physiol. 152, 169–182.

Ohguro, H., Rudnicka-Nawrot, M., Buczylko, J., Zhao, X., Taylor, J.A., Walsh, K.A., Palczowski, K., 1996. Structural and enzymatic aspects of rhodopsin phosphorylation. J. Biol. Chem. 271, 5215–5224.

O'Tousa, J.E., 1992. Requirement of N-linked glycosylation site in *Drosophila* rhodopsin. Vis. Neurosci. 8, 385–390.

Papatsenko, D., Sheng, G.J., Desplan, C., 1997. A new rhodopsin in R8 photoreceptors of *Drosophila* — evidence for coordinate expression with Rh3 in R7 cells. Development 124, 1665–1673.

Petrov, D., Losovskaya, E.R., Hartl, D., 1996. High intrinsic-rate of DNA loss in *Drosophila*. Nature 384, 346–349.

Rice, K., 1994. The origin, evolution, and classification of G-protein-coupled receptors. Ph.D. thesis, Harvard University.

Swofford, D.L., 1991. PAUP: Phylogenetic analysis using parsimony. Illinois Natural History Survey. In Progress in Sensory Physiology, Champaign, IL.

Thompson, J.D., Higgins, D.G., Gibson, T. J., 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Thurmond, R.L., Creuzenet, C., Reeves, P.J., Khorana, H.G., 1997. Structure and function in rhodopsin: peptide sequences in the cytoplasmic loops of rhodopsin are intimately involved in interaction with rhodopsin kinase. Proc. Natl. Acad. Sci. USA 94, 1715–1720.

Townson, S.M., Chang, B.S.W., Salcedo, E., Chadwell, L., Pierce, N.E., Britt, S.M., 1998. Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in *Drosophila*, and physiological characterization. J. Neurosci. 18, 2412–2422.

Turbeville, J.M., Staton, J.L., Brown, W.M., 1997. Amplification of the complete mitochondrial genome of two protostome worms: a useful technique for comparative studies of metazoan mitochondrial DNA. Mol. Mar. Biol. Biotechnol 6, 141–143.

Zuker, C.S., Mismer, D., Hardy, R., Rubin, G.M., 1988. Ectopic expression of a minor *Drosophila* opsin in the major photoreceptor cell class: distinguishing the role of primary receptor and cellular context. Cell 55, 475–482.