Molecular Characterization of Crustacean Visual Pigments and the Evolution of Pancrustacean Opsins

Megan L. Porter,* Thomas W. Cronin,† David A. McClellan,‡ and Keith A. Crandall‡

*Department of Microbiology and Molecular Biology, Brigham Young University; †Department of Biological Sciences, University of Maryland Baltimore County; and ‡Department of Integrative Biology, Brigham Young University

Investigations of opsin evolution outside of vertebrate systems have long been focused on insect visual pigments, whereas other groups have received little attention. Furthermore, few studies have explicitly investigated the selective influences across all the currently characterized arthropod opsins. In this study, we contribute to the knowledge of crustacean opsins by sequencing 1 opsin gene each from 6 previously uncharacterized crustacean species (*Euphausia superba*, *Homarus gammarus*, *Archaeomysis grebnitzkii*, *Holmesimysis costata*, *Mysis diluviana*, and *Neomysis americana*). Visual pigment spectral absorbances were measured using microspectrophotometry for species not previously characterized (*A. grebnitzkii* = 496 nm, *H. costata* = 512 nm, *M. diluviana* = 501 nm, and *N. americana* = 520 nm). These novel crustacean opsin sequences were included in a phylogenetic analysis with previously characterized arthropod opsin sequences to determine the evolutionary placement relative to the well-established insect spectral clades (long-/middle-/short-wavelength sensitive). Phylogenetic analyses indicate these novel crustacean opsins form a monophyletic clade with previously characterized crayfish opsin sequences and form a sister group to insect middle-/long-wavelength–sensitive opsins. The reconstructed opsin phylogeny and the corresponding spectral data for each sequence were used to investigate selective influences within arthropod, and mainly "pancrustacean," opsin evolution using standard d_N/d_S ratio methods and more sensitive techniques investigating the amino acid property changes resulting from nonsynonymous replacements in a historical (i.e., phylogenetic) context. Although the conservative d_N/d_S methods did not detect any selection, 4 amino acid properties (coil tendencies, compressibility, power to be at the middle of an α-helix, and refractive index) were found to be influenced by destabilizing positive selection. Ten amino acid sites relating to these properties were found to face the binding pocket, within 4 Å of the chromophore and thus have the potential to affect spectral tuning.

Introduction

Visual pigment research has long been of interest to a number of biological disciplines, including sensory ecologists, visual physiologists, biochemists, and molecular evolutionists. Composed of a chromophore bound to an integral membrane protein (opsin), visual pigments are phenotypically characterized by the wavelength of maximal absorption (λ_{max}). As most visual pigments contain the same chromophore, variation in opsin sequence is responsible for most of the observed variation in λ_{max}. The ability to study the genetic mechanisms behind phenotypic variation has made opsin a model evolutionary system. In vertebrates, these studies often take the form of site-directed mutagenesis studies, where the effect of changes at a single amino acid residue on λ_{max} can be quantified (Nathans 1990a; Neitz et al. 1991; Asenjo et al. 1994; Yokoyama and Radlwimmer 1999; Wilkie et al. 2000; Yokoyama et al. 2000; Yokoyama and Radlwimmer 2001; Cowing et al. 2002; Yokoyama and Tada 2003). This work suggests that spectral tuning involves the distribution of charged or polar residues within the chromophore-binding pocket (Nathans 1990a; Neitz et al. 1991; Chan et al. 1992). Due to the difficulty of expressing visual pigments, most studies of "nonvertebrates" (i.e., insects, cephalopods, crayfish, and *Limulus*) have used comparative methods and homology modeling to identify sites potentially important in spectral tuning (Chang et al. 1995; Chase et al. 1997; Crandall and Cronin 1997; Briscoe 2001, 2002), whereas a brave few have investigated spectral tuning using *Drosophila*-based heterologous expression systems (Britt et al. 1993;

Salcedo et al. 1999, 2003). These studies have not only uncovered some similarities among spectral tuning sites in vertebrates and insects (Briscoe 2001; Salcedo et al. 2003), but also illustrated that there are a number of unique residues affecting spectral tuning in insects.

Given that the visual pigment genes diversified into the observed spectral clades sometime after the separation of the protostomes and deuterostomes (Pichaud et al. 1999; Arendt et al. 2004), some differences in functionality are expected. In fact, several significant differences in the mechanism of photoactivation have been documented, most notably differences related to the Schiff base. In vertebrates, a negatively charged counterion (E113) stabilizes the positive charge of the Schiff base (Sakmar et al. 1989; Nathans 1990b; Zhukovsky et al. 1992). However, in insects and cephalopods, this same site is a tyrosine (or in UV-sensitive opsins, a phenylalanine). Studies have shown that this residue does not function as a counterion, and the exact method of photoactivation is not yet completely understood (Nakagawa et al. 1999; Salcedo et al. 2003).

Although a plethora of studies have looked at structure/function relationships in vertebrate opsins from an evolutionary perspective, few studies have investigated similar issues in the arthropod visual pigment diversification. Consequently, much less is known about the mode and tempo of opsin evolution in these animals and most of the evolutionary work related to visual pigments has focused on insect systems. Insects have at least 4 main spectral classes: long-wavelength–sensitive (LWS), middle-wavelength–sensitive (MWS), and 2 short-wavelength–sensitive (SWS) groups (UV and blue) (O'Tousa et al. 1985; Montell et al. 1987; Zuker et al. 1987; Feiler et al. 1988; Carulli and Hartl 1992; Feiler et al. 1992; Carulli et al. 1994; Towner and Gartner 1994; Smith et al. 1997; Salcedo et al. 1999; Briscoe 2001, 2002; Briscoe and Chittka 2001; Taylor et al. 2005). Although there have been a large number of physiological

studies of spectral sensitivities in other major arthropod groups, particularly within crustaceans, how these spectral variants relate to the defined insect opsin clades is unknown. The only other nonvertebrate taxa in which opsin sequences have been explicitly investigated are horseshoe crabs (Smith et al. 1993), molluscs (Brown PK and Brown PS 1958; Hubbard and St George 1958; Hall et al. 1991; Morris et al. 1993), and crustaceans (Sakamoto et al. 1996; Crandall and Cronin 1997; Crandall and Hillis 1997; Oakley and Huber 2004).

Here we investigate opsin evolution in insect, crustacean, and mollusc systems and add to the knowledge of arthropod opsins by expanding the characterized opsin sequences from within the Crustacea. Although a number of insect groups have been the focus of evolutionary studies (most notably *Drosophila* and Lepidoptera; Carulli et al. 1994; Briscoe 2000, 2001, 2002) and representatives from known insect spectral variants are often used to place phylogenetically novel opsins and to form hypotheses about uncharacterized spectral sensitivities, few studies have focused on explicitly investigating the selective forces acting upon this diversity of opsins. We have identified opsin sequences from 6 previously uncharacterized crustacean species and determined the λ_{max} in those species without previous estimates. These data complement previously characterized insect opsins, expanding the data set available to investigate selective influences from insects to "pancrustacea." Furthermore, within these data are many opsin sequences with corresponding characterization of the visual pigment spectral sensitivities. By combing the molecular and physiological data sets and analyzing them with established (PAML; Yang 1997) and newly described (TreeSAAP; Woolley et al. 2003) methods for detecting selection, we test the hypothesis that visual pigment spectral tuning throughout animal diversity is related to the distribution of charged or polar residues relative to the chromophore-binding pocket (Nathans 1990a; Neitz et al. 1991; Chan et al. 1992; Salcedo et al. 2003).

Methods
Taxon Sampling and Outgroup Choice

Opsin sequence data were collected from 6 crustacean species: *Archaeomysis grebnitzkii*, *Holmesimysis costata*, *Mysis diluviana*, *Neomysis americana* (Mysida); *Homarus gammarus* (Decapoda); and *Euphausia superba* (Euphausiacea). In addition, 3 sequences from the stomatopod crustacean *Neogonodactylus oerstedii* that were characterized for a Ph.D. thesis (Brown 1996) were included in our analyses. To these data, we added any opsin sequence data for nonvertebrate species that represented at least half of the transmembrane-spanning portion of the gene and had corresponding λ_{max} values previously reported in the literature (table 1). Sequences chosen for analyses were limited to those with spectral characterization in order to evaluate any sites identified to be under selection in terms of spectral tuning of the visual pigments. Outgroups were chosen from the vertebrate opsin clade, based on the hypothesis that duplication of opsin genes occurred independently in the lines of descent leading to protostomes and deuterostomes (Pichaud et al. 1999). Bovine rhodopsin was included as the only G protein–coupled receptor (GPCR) to have been

crystallized (Palczewski et al. 2000). The inclusion of this sequence in our alignment allowed for sites identified by selection detection methods to be mapped to the bovine protein structure (see below). Two representatives of pineal opsin and an "orphan" human opsin paralog (GPR52) were included as outgroups, representing basal lineages to the vertebrate visual pigment clade (Bellingham et al. 2003; Fredriksson et al. 2003). Finally, human melatonin receptor 1A, a GPCR that has been phylogenetically placed close to the human visual pigment clade (Fredriksson et al. 2003), was chosen as a distant outgroup.

Microspectrophotometry

Of the 6 crustacean species molecularly characterized, 2 have been previously characterized with respect to spectral sensitivities: *H. gammarus* and *E. superba*. In order to link the analyses of selection at the molecular level to function (i.e., spectral tuning), spectral sensitivities of the 4 remaining species were characterized using microspectrophotometry (MSP). For spectral analyses, live specimens were shipped overnight in dark conditions to the University of Maryland Baltimore County. In most cases, animals were used within a week of collection. All organisms were dark adapted at least overnight, but more commonly for several days, before use. Eyes were removed under dim red light, mounted in tissue medium, and flash frozen. Frozen eye samples were sectioned immediately using a cryostat microtome to produce ~14-μm thick sections. Individual sections were mounted on cover slips and scanned under dim red light on a microscope for usable rhabdom structures. Suitable sections were mounted in Ringer's buffer solution between coverslips sealed with a ring of silicone grease. The equipment and general procedure used for MSP have been described by Cronin et al. (1996). Briefly, a linearly polarized scanning beam was placed within a single rhabdom. Scans were made from 400 to 700 nm, with measurements taken at 1-nm steps. Each dark-adapted rhabdom was scanned twice to check for stability. If the 2 scans were identical, the second was saved as the direct absorption spectrum of the dark-adapted photoreceptor. The rhabdom was then exposed to 2 minutes of bright white light, followed by a second absorption scan. During bright-light treatments, the field diaphragm of the substage illuminator was closed down to produce a spot of ~10 μm diameter at the level of the rhabdom, minimizing local heating of the preparation. For each photoreceptor, the rhodopsin absorption spectrum was taken as the difference between the initial, dark-adapted spectrum and the final, photobleached spectrum. The wavelength of maximum absorption (λ_{max}) was estimated for each difference spectrum using a least squares procedure (see Cronin et al. 1994). We compared all photobleach difference spectra with standard rhodopsin and porphyropsin templates derived by Stavenga et al. (1993) and subsequently averaged together those that closely resembled either template. The average spectrum was then fitted to the corresponding template again to determine a λ_{max} value that best represents the spectra of the measured visual pigments. Results from 6 to 10 rhabdoms, representing 2 or more individuals, were obtained for each species.

Table 1
Taxonomy, GenBank Accession Numbers (http://www.ncbi.nlm.nih.gov/) for Gene Sequences and Wavelength of Maximal Absorbance (λ_{max}) and References for Opsins Analyzed in This Study. For λ_{max} Values, "m" Denotes Measurements from Males and "f" from Females, Whereas "*" Indicates Averaged Values. Data New to This Study Are Indicated in Bold

Taxon	Accession #	λ_{max} (nm)	λ_{max} Reference
Mollusca Rh			
Cephalopoda			
Loligo forbesi	X56788	494	Morris et al. (1993)
Loligo pealii	AY450853	493	Brown PK and Brown PS (1958), Hubbard and St George (1958)
Loligo subulata	Z49108	499	Morris et al. (1993)
Sepia officinalis	AF000947	492	Brown PK and Brown PS (1958)
Todarodes pacificus	X70498	480	Naito et al. (1981)
Enteroctopus dofleini	X07797	475	Koutalos et al. (1989)
Arthropoda LWS			
Chelicerata			
Limulus polyphemus (lateral eye)	L03781	520	Hubbard and Wald (1960)
L. polyphemus (ocelli)	L03782	530	Nolte and Brown (1972)
Crustacea			
Euphausia superba	**DQ852576–DQ852580**	487	Frank and Widder (1999)
Homarus gammarus	**DQ852587–DQ852590**	515	Kent (1997)
Cambarellus shufeldtii	AF003544	526	Crandall and Cronin (1997)
Cambarus ludovicianus	AF003543	529	Crandall and Cronin (1997)
Orconectes virilis	AF003545	530	Goldsmith (1978), Cronin and Goldsmith (1982)
Procambarus milleri	AF003546	522	Cronin and Goldsmith (1982), Crandall and Cronin (1997)
Procambarus clarkii	S53494	533	Zeiger and Goldsmith (1994)
Archaeomysis grebnitzkii	**DQ852573–DQ852575**	**496**	**This study**
Holmesimysis costata	**DQ852581–DQ852586**	**512**	**This study**
Mysis diluviana	**DQ852591**	**501**	**This study**
Neomysis americana	**DQ852592–DQ852598**	**520**	**This study**
Neogonodactylus oerstedii Rh1	**DQ646869**	489	Cronin and Marshall (1989)
N. oerstedii Rh2	**DQ646870**	528	Cronin and Marshall (1989)
N. oerstedii Rh3	**DQ646871**	522	Cronin and Marshall (1989)
Insecta			
Manduca sexta	L78080	520	White et al. (1983)
Spodoptera exigua	AF385331	515	Langer et al. (1979)
Galleria mellonella	AF385330	510	Goldman et al. (1975)
Papilio xuthus Rh1	AB007423	520	Arikawa et al. (1987, 1999), Kitamoto et al. (1998)
P. xuthus Rh2	AB007424	520	Arikawa et al. (1987, 1999), Kitamoto et al. (1998)
P. xuthus Rh3	AB007425	575	Arikawa et al. (1987, 1999), Kitamoto et al. (1998)
Pieris rapae	AB177984	540	Ichikawa and Tateda (1982)
Vanessa cardui	AF385333	530	Briscoe et al. (2003)
Junonia coenia	AF385332	510	Briscoe (2001)
Heliconius erato	AF126750	570	Struwe (1972)
Heliconius sara	AF126753	550	Struwe (1972)
Bicyclus anynana	AF484249	560	Vanhoutte et al. (2002)
Camponotus abdominalis	U32502	510	Popp et al. (1996)
Cataglyphis bombycinus	U32501	510	Popp et al. (1996)
Apis mellifera	U26026	529 m, 540 f	Peitsch et al. (1992)
Bombus terrestris	AY485301	529	Peitsch et al. (1992)
Osmia rufa	AY572828	553	Peitsch et al. (1992)
Schistocerca gregaria	X80071	520	Gartner and Towner (1995)
Sphodromantis sp.	X71665	515*	Rossel (1979)
Arthropoda MWS			
Crustacea			
Hemigrapsus sanguineus	D50583, D50584	480	Sakamoto et al. (1996)
Insecta			
Drosophila melanogaster Rh6	Z86118	508	Salcedo et al. (1999)
D. melanogaster Rh1	AH001026	478	Feiler et al. (1988)
Calliphora erythrocephala Rh1	M58334	490	Paul et al. (1986)
D. melanogaster Rh2	M12896	420	Feiler et al. (1988)
Insecta blue			
S. gregaria	X80072	430	Gartner and Towner (1995)
M. sexta	AD001674	450	White et al. (1983)
P. xuthus Rh4	AB028217	460	Eguchi et al. (1982), Arikawa et al. (1987)

Table 1
Continued

Taxon	Accession #	λ_{max} (nm)	λ_{max} Reference
A. mellifera	AF004168	439	Townson et al. (1998)
D. melanogaster Rh5	U67905	437	Salcedo et al. (1999)
Insecta UV			
A. mellifera	AF004169	353	Townson et al. (1998)
C. abdominalis	AF042788	360	Smith et al. (1997)
C. bombycinus	AF042787	360	Smith et al. (1997)
M. sexta	L78081	357	White et al. (1983)
P. xuthus Rh5	AB028218	—	—
D. melanogaster Rh4	AH001040	375	Feiler et al. (1992)
D. melanogaster Rh3	M17718	345	Feiler et al. (1992)
Outgroups			
Bos taurus rhodopsin	AH001149	—	—
Gallus gallus pineal opsin	U15762	—	—
Anolis carolinensis pineal opsin	AH007737	—	—
Homo sapiens GPR52	NM_005684	—	—
H. sapiens melatonin receptor 1A	NM_005958	—	—

DNA Extraction, Polymerase Chain Reaction, Cloning, and Sequencing

All specimens were stored in 70–95% ethanol and kept at 4 °C until extracted. Genomic DNA was extracted using DNeasy kits (Qiagen) following the manufacturer's instructions. Polymerase chain reaction (PCR, Saiki et al. 1988) products for the opsin gene were amplified using a seminested degenerate PCR strategy. An initial PCR using primers LWF1a: 5′-TGG TAY CAR TWY CCI CCI ATG AA-3′ and OPSRD: 5′-CCR TAN ACR ATN GGR TTR TA-3′ (Chang et al. 1996) with standard conditions (2.5 µl 10× *Taq* buffer, 4 µl 10 mM deoxynucleoside triphosphates, 2.5 µl each of two 10 mM primers, 1.25 U Eppendorf HotMaster *Taq*, ∼12.5 µl double distilled water) was run on a Perkin-Elmer 9700 machine for 35 cycles of 95 °C for 30 s, 48 °C for 45 s, and 70 °C for 1:15 min, followed by chain extension at 72 °C for 15 min. The first round PCR reactions were then diluted 1:10 with sterile water and used as template for a second round PCR of another 35 cycles using primers LWF1a and Scylla: 5′-TTR TAI ACI GCR TTI GCY TTI GCR AA-3′ (Taylor et al. 2005). Second round PCR products were visualized by agarose (1.2%) gel electrophoresis. Visible opsin DNA bands were excised and cleaned from the agarose gel using a GeneClean II kit (Bio 101). Purified opsin PCR products were cloned using the TOPO TA Cloning Kit (Invitrogen, San Diego, CA), following manufacturer's instructions. For each species examined, 3–10 clones were lysed in 50 µl buffer (10 mM Tris-HCl; 0.1 mM ethylenediaminetetraacetic acid, pH 8.0) for 10 min at 96 °C and inserts were PCR amplified from lysed cells using plasmid-specific primer pairs M13(-20) (5′-GTAAAACGACGGCCAGT-3′) and M13(-24) (5′-AACAGCTATGACCATG-3′) and the following PCR conditions: denaturation at 94 °C for 1 min, primer annealing at 55 °C for 1 min, and chain extension at 72 °C for 3 min, for 30 cycles. Cloned PCR products were purified using the Millipore Montage purification system and sequenced in both directions on an ABI Prism 3730 capillary autosequencer using the ABI big-dye Ready-Reaction kit and following the standard cycle sequencing protocol but using 1/16th of the suggested reaction volume.

Phylogenetic Analyses

Opsin sequences were either generated in the lab (see above section) or were downloaded from GenBank (http://www.ncbi.nlm.nih.gov; see table 1). All of the species included in the analyses were from species where both the opsin sequence and the λ_{max} have been previously characterized or were measured in this study. Only sequences that spanned more than half of the opsin transmembrane domains were used to reduce spurious arrangements caused by short sequences and to increase the power of tests of selection by decreasing missing data (see below section). MacClade (Maddison DR and Maddison WP 2001) was used to create an initial nucleotide sequence alignment, which was then converted to amino acid sequences for alignment. Amino acid sequences were aligned in ClustalX (Thompson et al. 1997) using a Blosum log2 weight matrix generated using MatrixGen (http://matrixgen.sourceforge.net/) with aligned class a GPCRs from the GPCR Web site (http://www.gpcr.org). The initial nucleotide sequence alignment was then adjusted to match the resulting amino acid alignment and fine-tuned based on structural information (Chang et al. 1995; Palczewski et al. 2000). The final nucleotide alignment is available online (http://inbio.byu.edu/faculty/kac/crandall_lab/pubs.html).

Phylogenetic trees were reconstructed from both the nucleotide and amino acid alignments. The nucleotide phylogeny was reconstructed using Bayesian methods coupled with Markov chain Monte Carlo (BMCMC) inference as implemented in MrBayes v3.04b (Ronquist and Huelsenbeck 2003). Because different codon positions have different functional constraints, the data set was partitioned into first-, second-, and third-codon positions for mixed-model analyses. Model selection for each partition followed the procedure outlined by Posada and Buckley (2004) for Akaike Information Criterion (AIC) as implemented in Modeltest v3.6 (Posada and Crandall 1998). Four independent BMCMC analyses were run with each consisting of 4

chains. Each Markov chain was started from a random tree and run for 6.0×10^6 cycles, sampling every 1,000th generation. Model parameters were treated as unknown variables with uniform default priors and were estimated as part of the analysis. To confirm that our Bayesian analyses converged and mixed well, we monitored the fluctuating value of likelihood and all phylogenetic parameters graphically and compared means and variances of all likelihood parameters and likelihood scores from independent runs using the program Tracer v1.2 (Rambaut and Drummond 2003). All sample points prior to reaching stationary were discarded as burn-in. The posterior probabilities (pP) for individual clades obtained from separate analyses were compared for congruence and then combined and summarized on a majority-rule consensus tree (Huelsenbeck and Imennov 2002; Huelsenbeck et al. 2002). A phylogeny based on the amino acid alignment was constructed in the program PHYML (Guindon and Gascuel 2003), which allows for the fast estimation of large data sets within a maximum likelihood (ML) framework. The best-fit model for the amino acid alignment was determined using ProtTest v1.2.6 (Abascal et al. 2005), which uses the phylogenetic analyses library in conjunction with PHYML to compute the likelihood for each of 64 candidate models of protein evolution. The fit of each of these models to our data set was then determined using a second-order AIC_c framework with sample size equal to the total number of characters (i.e., alignment length). Branch support values were estimated from 100 PHYML bootstrap replicates as bootstrap proportions (BP).

Investigating Selective Influences

The influence of selective forces on the evolution of visual pigments was investigated using a suite of methodologies. First, selection was investigated using nonsynonymous to synonymous substitution rate ratios (d_N/d_S) as calculated in a likelihood framework in the CODEML module of PAML v3.14 (Yang 1997). Selection was determined by using likelihood ratio tests to evaluate nested site-specific models with and without incorporating selection (Yang et al. 2000). For site-specific models, models M1a (nearly neutral) versus M2a (positive selection) and M7 (beta) versus M8 (beta&ω) were tested as suggested in the PAML v3.14 manual. For site classes where $\omega > 1$, Bayes empirical Bayes calculations of posterior probabilities are implemented in models M2a and M8 to identify the particular sites under positive selection (Yang et al. 2005). All models were run twice with starting omega values of less than and greater than 1 to test for entrapment in local optima (Yang et al. 2000). Likelihood ratio tests, to determine whether particular models provided a significantly better fit to the data, were performed by comparing the likelihood ratio test statistic ($-2[\ln L1 - \ln L2]$) with critical values of the Chi square distribution with the appropriate degrees of freedom (Yang 1997).

A common measure of selective pressure in protein-coding genes is $d_N/d_S = \omega$, where $\omega > 1$, $=1$, or <1 indicate positive selection, neutral evolution, and purifying selection, respectively. Many advances have been made in the estimation of this ratio that increase its power of detecting selection, including models that are either lineage or site specific (Nielsen and Yang 1998; Yang 1998; Yang and Nielsen 1998; Yang et al. 2000; Anisimova et al. 2002; Forsberg and Christiansen 2003; Bielawski and Yang 2004). Site-specific models will be used to investigate the presence of positive selection across the entire pancrustacean opsin phylogeny. However, if adaptive evolution occurs at only a few time points and affects only a few amino acids, site-specific models may still lack power in detecting positive selection (McClellan et al. 2005). For protein-coding genes such as opsin where strong structural and functional constraints can lead to a large proportion of invariable residues (31–99% similarity in arthropod opsin sequences), d_N/d_S methods are still likely to be very conservative. Furthermore, beyond detecting the presence of positive selection at either specific sites or lineages within a phylogeny, d_N/d_S methods do not provide information on the type of positive selection detected (directional or nondirectional, stabilizing or destabilizing), have very little power to detect purifying selection, and offer little insight into how the identified selection affects the overall structure and function of the protein.

Recent methods have taken the investigation of selection in protein-coding genes further by addressing several of these issues. A different approach for detecting selection in amino acid sequences is to look at the magnitudes of property change of nonsynonymous residues across a phylogeny. Amino acid substitutions have a wide range of effects on a protein depending on the difference in physicochemical properties and location in the protein structure. This approach provides further resolution to differentiating between types of selective pressures with the ability to detect positive and negative and stabilizing and destabilizing selection and offers insights into the structural and functional consequences of the identified residues under selection (McClellan et al. 2005). We used the program TreeSAAP v3.2 (Woolley et al. 2003) to test for selection on amino acid properties within our opsin data set. For each property examined, a range of possible 1-step changes as governed by the structure of the genetic code was determined and divided into 8 magnitude categories of equal range, with lower categories indicating more conservative changes and higher categories denoting more radical changes. In order to construct an expected distribution of amino acid property change, each of the 9-nt changes in every codon of every DNA sequence within the data set was evaluated, with each nonsynonymous change assigned to 1 of the magnitude categories for each property independently. These property changes were then summed across the data set, constructing a set of relative frequencies of change for each of the 8 magnitude categories to establish the null hypothesis under the assumption of neutral conditions (McClellan and McCracken 2001). If distributions of observed changes fail to fit the expected distributions based on goodness-of-fit scores and z-scores, the null hypothesis of neutrality is rejected. Because we are interested in the evolution of spectral variation, we targeted sites identified to be under positive destabilizing selection, defined as selection for radical amino acid changes resulting in structural or functional shifts in local regions of the protein (McClellan et al. 2005). In terms of TreeSAAP analyses, positive destabilizing selection is
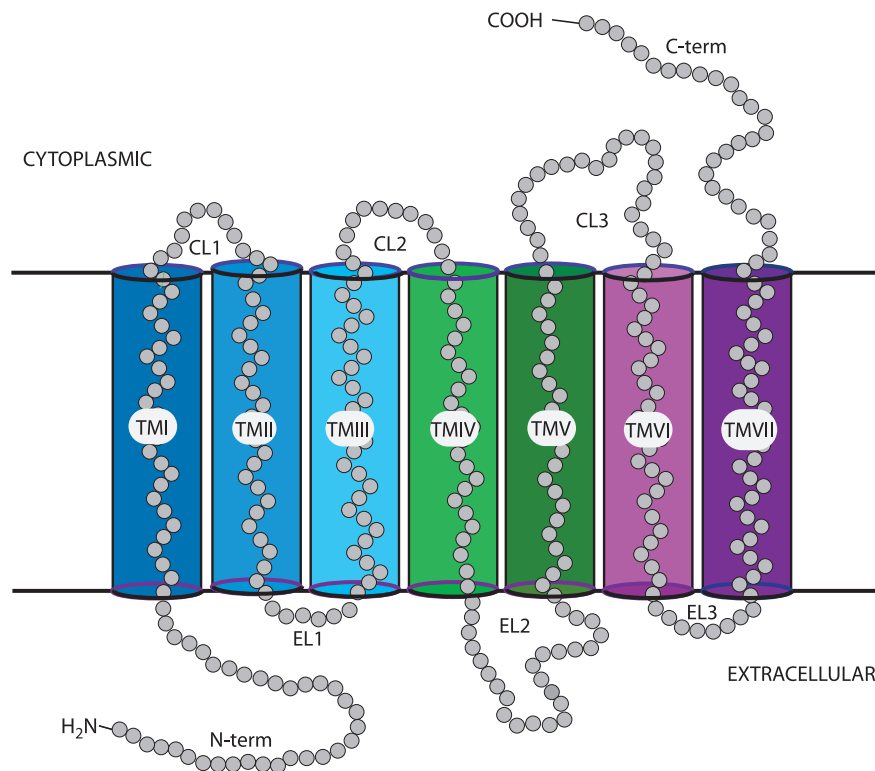
FIG. 1.—A 2-dimensional schematic of an opsin protein, with domains labeled as they are referred to in the paper. Opsin is composed of 3 general domains: extracellular, transmembrane, and cytoplasmic. In common with all known GPCRs, the opsin protein contains 7 transmembrane–spanning α-helices, here labeled TMI–TMVII. The extracellular domain contains the N terminus (N-term) and 3 interhelix loops (EL1, EL2, and EL3). Similarly, the cytoplasmic domain includes 3 loops (CL1, CL2, and CL3) and the C terminus (C-term). Other structural features of the opsin protein (i.e., beta sheets and the eighth cytoplasmic α-helix) are not discussed in this study and therefore are not illustrated here.

defined a priori as properties with significantly greater amino acid replacements than neutral expectations for magnitude categories 7 and 8 (e.g., the 2 most radical property change categories). Using this criterion, identified sites were mapped onto the protein structure to investigate the structural/functional impact of the selection. Within Tree-SAAP, 31 amino acid properties are evaluated across a phylogeny using either the entire data set or a sliding window analysis. For our purposes, properties and magnitudes showing significantly more observed than expected numbers of changes at the $P \leq 0.001$ level were first identified with an overall analysis of our data. The identified properties were then subjected to a sliding window analysis, investigating varying window sizes (15, 20, 30, and 40 codons in width) to determine the range that maximizes the signal. The results of the sliding window analyses were used to identify regions in the protein that differ significantly from a nearly neutral model at a significance level of $P = 0.001$. Finally, we identified the particular amino acid residues within each of these regions that contained positive destabilizing selection for each property. Using a high-resolution (2.6 Å) bovine rhodopsin template (1L9H.pdb) from the Protein Data Bank (http://www.rcsb.org/pdb/; Berman et al. 2000) with the program Swiss-PdbViewer v.3.7 (http://www.expasy.org/spdbv; Guex and Peitsch 1997), we mapped the identified residues to the opsin protein structure using our alignment as a template. Using features of this program, residues were evaluated for their proximity to the chromophore-binding

pocket using a conservative distance of 4 Å to infer a residue's potential interaction with the chromophore. Previous studies of functionally important amino acids in class A GPCRs demonstrated that there are 2 classes of residues: those that mediate conformational change and receptor activation of G proteins, which are expected to be similar across receptors, and those that mediate ligand binding, which are expected to be highly specific due to the structural diversity of ligands (Madabushi et al. 2004). By limiting our selection of identified amino acids to those with potential to interact with the chromophore, we are looking at those residues most likely to be important to opsin-specific functioning rather than of general importance to GPCR function. Throughout this paper, domains are labeled as indicated in figure 1. All references to amino acid residues are given using bovine rhodopsin numbering to make inferences directly comparable to previous studies.

Results

MSP and Molecular Characterization of Novel Crustacean Opsin Sequences

The mysid visual pigment λ_{max} ranged from 496 to 520 nm (fig. 2). Most rhabdom scans fit the Stavenga et al. (1993) rhodopsin template, but several scans from *M. diluviana* indicated the presence of a porphyropsin pigment (utilizing the A2 chromophore). Here, we present only the rhodopsin data; the porphyropsin data are presented elsewhere (Jokela-Määttä et al. 2005). Besides the *M. diluviana*
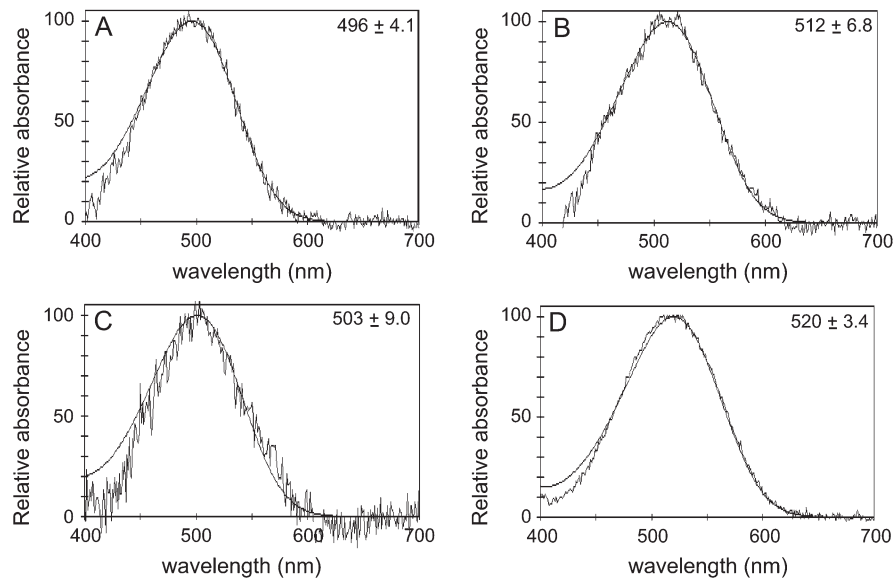
Fig. 2.—Average difference spectra for photobleaching of rhabdoms in retinal sections from 4 mysid species: (*A*) *Archaeomysis grebnitzkii*, (*B*) *Holmesimysis costata*, (*C*) *Mysis diluviana*, and (*D*) *Neomysis americana*. Each panel displays the results from the indicated species together with the spectrum of the best-fit rhodopsin template (Stavenga et al. 1993). The average wavelength of absorption ± standard deviation is indicated in each panel. The numbers of rhabdoms included in each average spectrum are as follows: *A. grebnitzkii*—12; *H. costata*—10; *M. diluviana*—6; and *N. americana*—13.

porphyropsin, there was no evidence for the presence of additional spectral classes of visual pigments in any of the species investigated.

The isolated crustacean opsin sequences were 284 amino acids in length and spanned from 9 amino acids before the start of transmembrane α-helix (TM) I to the middle of TMVII (AA site 299). Sequences were submitted to GenBank under the accession numbers: DQ852573–DQ852598 (newly characterized crustacean species) and DQ646869–DQ646871 (*N. oerstedii*). Similar to genes previously sequenced from crayfish and cephalopods (Morris et al. 1993; Crandall and Cronin 1997; Crandall and Hillis 1997), none of the isolated crustacean sequences contained any introns. Within each species, the clone sequence variability was less than 1%. This variability is less than expected from *Taq* error (2.3×10^{-6} to 2.1×10^{-4} errors per basepair per cycle; Bracho et al. 1998) given our PCR methodologies (2 rounds of initial PCR, cloning, and PCR confirmation of cloned products), and therefore clones were considered to be identical and were merged into a consensus sequence for each species for further analyses. When aligned with other opsin sequences, the new crustacean sequences exhibited the characteristic indel region of arthropods in cytoplasmic loop (CL) 3 of ~14 amino acids.

Phylogenetic Analyses of Pancrustacean Opsins

For each codon position, the best-fit models all corresponded to general-time reversible model (GTR) + I + G (AIC$_w$: position 1 = 0.56; position 2 = 0.92; position 3 = 0.90). For the amino acid data, the best-fit model was WAG + G + F (AIC$_{cw}$ = 0.77, α = 1.1) (Whelan and Goldman 2001). Phylogenetic analyses using these models with either nucleotide or amino acid data produced similar trees, with only minor differences in the relationships of tip taxa

within well-supported clades. Both sets of analyses placed the novel crustacean opsin sequences with previously characterized crustacean sequences in a monophyletic clade sister to the insect MWS/LWS clade (fig. 3). Within each insect spectral clade, species cluster roughly by higher-order taxonomy (i.e., Lepidoptera, Hymenoptera, and Diptera). In contrast, although the crustacean clade is strongly supported (BP = 98; p*P* = 1.00), the currently sampled taxa exhibit very little taxonomic clustering with generally low branch support values. For example, neither the decapod species (crayfish lineage + *H. gammarus*) nor the 4 mysid species form monophyletic groups. The 3 genes sequenced from the stomatopod *N. oerstedii* do cluster, although branch support is low for the placement of *N. oerstedii* Rh2 (BP < 50, p*P* = 0.89). Interestingly, the *N. oerstedii* opsins do not cluster by spectral sensitivity, with strong support for the Rh1 (the corresponding visual pigment has λ$_{max}$ = 489 nm; Cronin and Marshall 1989) + Rh3 (λ$_{max}$ = 522 nm; Cronin and Marshall 1989) clade (BP = 100, p*P* = 1.00). Furthermore, the opsin located in the *N. oerstedii* peripheral ommatidia (Rh2) that is most similar to other crustacean compound eyes contains the most divergent opsin sequence (40.9–48.3% amino acid difference) relative to other crustacean LWS opsins.

The general topology of this tree demonstrates the presence of a monophyletic arthropod LWS clade containing representatives from the Hexapoda, Crustacea, and Chelicerata (BP = 100; p*P* = 1.00). Curiously, within the Diptera, there has been a gene duplication event leading to the formation of a distinct MWS clade, representatives of which have not yet been found in other insects (Spaethe and Briscoe 2004). A similar situation exists within the Crustacea, with the only sequenced MWS opsin (*Hemigrapsus sanguineus*; Sakamoto et al. 1996) falling outside of the main arthropod LWS clade, indicating an earlier gene duplication event; however, sequence data from other crustaceans
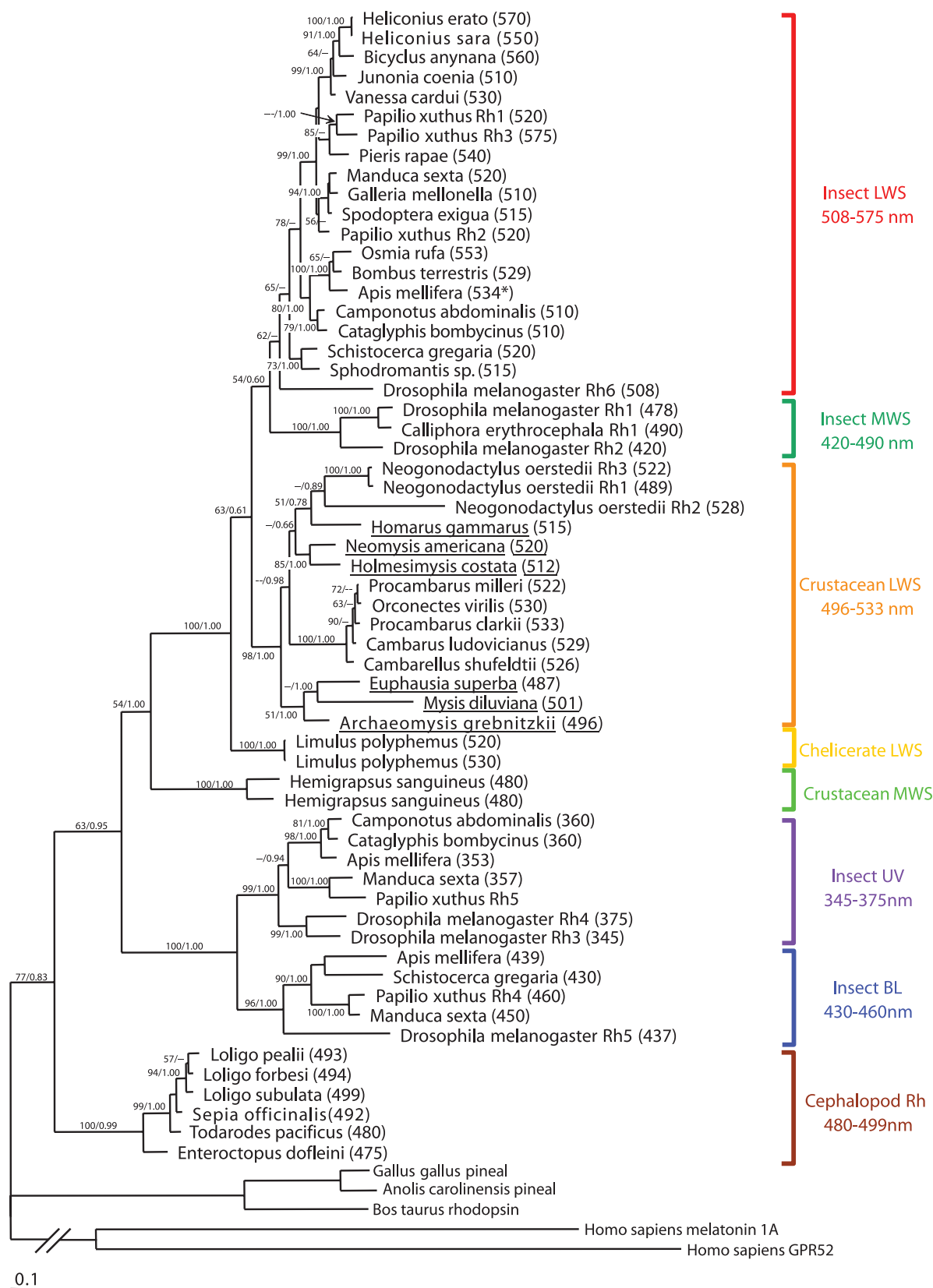
FIG. 3.—Phylogeny of arthropod and cephalopod opsins based on ML analyses of 304 amino acid residues. Numbers above or below each branch indicate ML BP/Bayesian pP. The λ_{max} for each taxon is given in parentheses after the species name. Spectral clades, with the λ_{max} variation for the represented taxa, are delineated. Accession numbers and references for λ_{max} values are given in table 1. New sequences and λ_{max} from this study are underlined.
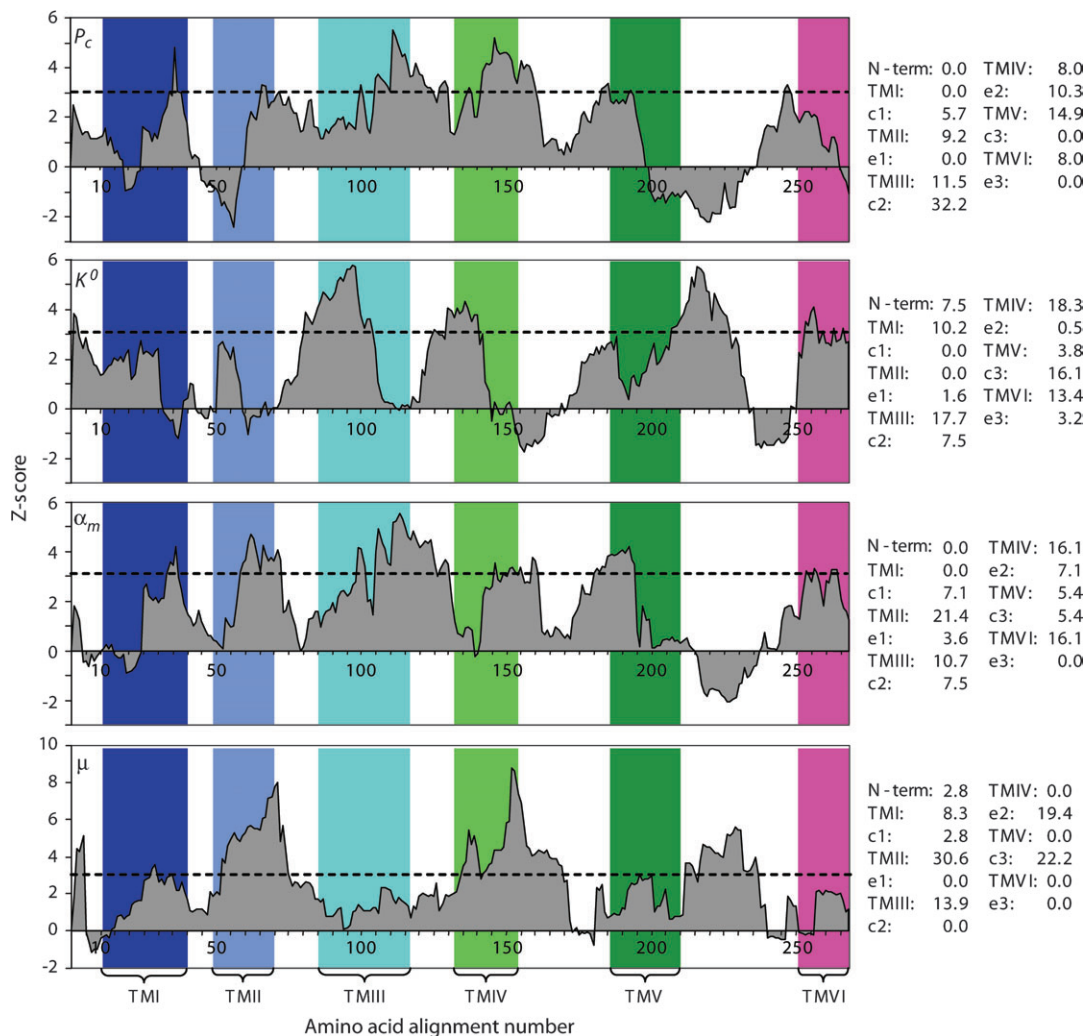
FIG. 4.—Results of TreeSAAP sliding window analyses using a window size of 20 and a step size of 1 amino acid for the 4 amino acid properties detected to be under positive destabilizing selection: coil tendencies (P_c), compressibility (K^0), power to be at the middle of the α-helix (α_m), and refractive index (μ). The dotted line in each panel indicates significance at the $P = 0.001$ level, with peaks above the line demonstrating regions of the protein under selection; areas under the line are not significantly different from neutral expectations. The transmembrane domains (I–VII) are indicated in each panel by colored boxes, with colors corresponding to figure 1. To the right of each panel is the percentage of sites for each property found in the individual domains.

are necessary to confirm that this gene is not a copy unique to *H. sanguineus*. Sister to the arthropod LWS/MWS is a monophyletic insect SWS clade (BP = 100; p*P* = 1.00), containing well-supported lineages representing both blue (BP = 96, p*P* = 1.00) and UV (BP = 99, p*P* = 1.00) spectral variants. Finally, sister to the arthropod visual pigment clade is the cephalopod lineage (BP = 100, p*P* = 0.99).

Selective Influences in Pancrustacean Opsins

Analyses of d_N/d_S ratios using PAML did not detect any sites under positive selection, and none of the models incorporating selection parameters contained site classes with ω > 1 or were significantly better than neutral models based on likelihood ratio tests. In contrast, TreeSAAP analyses identified 4 amino acid properties to be under positive destabilizing selection in our opsin data set ($P < 0.001$): coil tendencies, compressibility, power to be at the middle of the α-helix, and refractive index. Evaluation of various sliding window sizes indicated that a window of 20 amino

acids provided the best signal-to-noise ratio. Sliding window analyses using a 20 amino acid window size identified that selection on these properties occurred primarily in the transmembrane domains (fig. 4); however, the distribution of sites under selection within each domain differed between properties. For example, the majority of sites across the phylogeny exhibiting selection for refractive index were found in TMII (30.6%), TMIII (13.9%), CL3 (22.2%), and extracellular loop (EL) 2 (19.4%), whereas those found for coil tendencies were in CL2 (32.3%), TMV (14.9%), TMIII (11.5%), and EL2 (10.3%).

Using a high-resolution model of bovine rhodopsin (1LH9.pdb), the sites identified to be in regions of the protein under destabilizing selection were mapped to opsin (fig. 5). For compressibility, sites were concentrated on the extracellular end of TMI and in CL2 and scattered throughout TMIII, TMIV, and TMVI. Interestingly, every single amino acid change under selection identified for compressibility was a change to an alanine. Outside of the transmembrane domains, coil tendencies and power
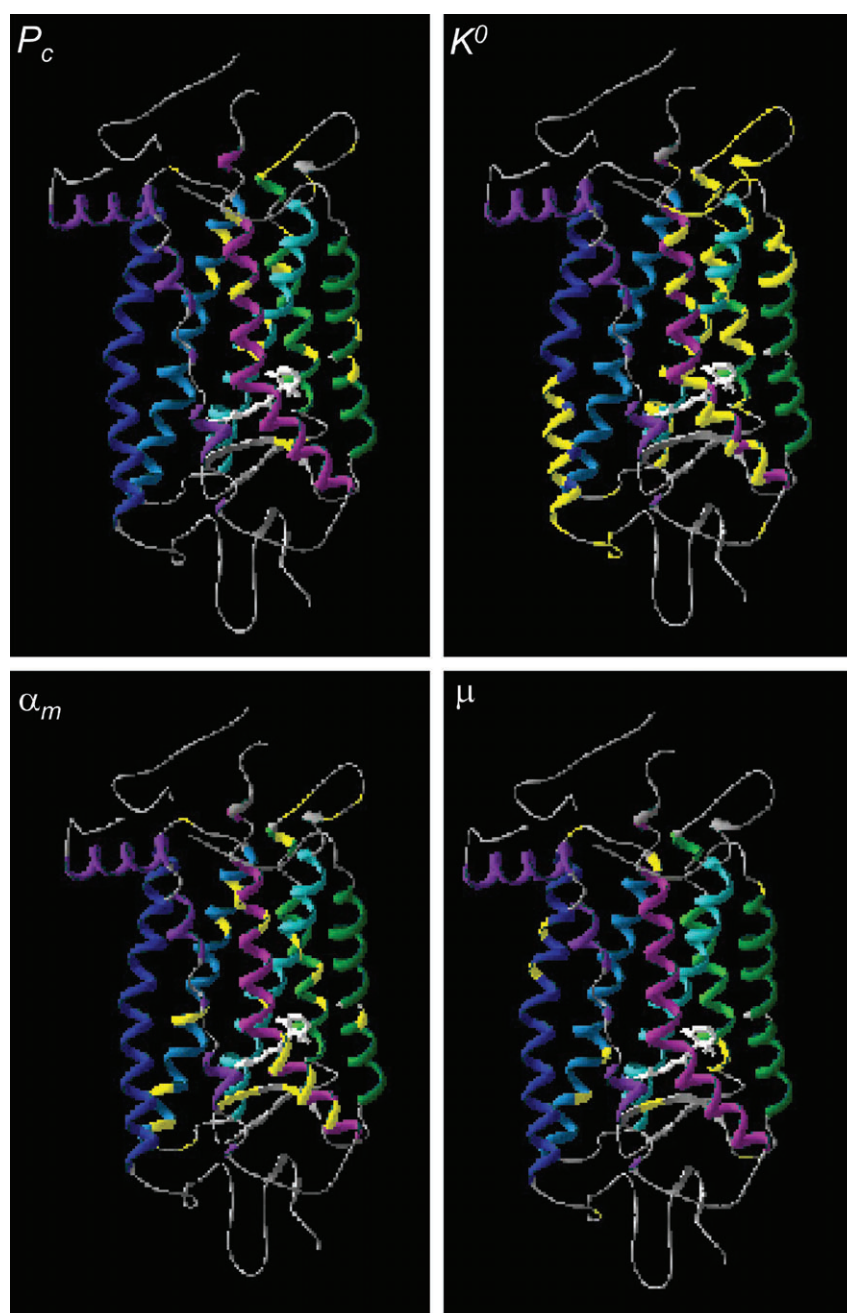
Fɪɢ. 5.—Amino acid residues identified by sliding window analyses for each property mapped to the high-resolution (2.6 Å) bovine rhodopsin structure. Residues with evidence for positive destabilizing selection are shaded yellow. The transmembrane domains are colored as in figure 1: TMI = dark blue, TMII = mid blue, TMIII = light blue, TMIV = light green, TMV = dark green, TMVI = light purple, and TMVII = dark purple. The chromophore in each panel is rendered in white. Each panel represents the sites identified for 1 property: coil tendencies (P_c), compressibility (K^0), power to be at the middle of the α-helix (α_m), and refractive index (μ).

to be at the middle of the α-helix also identified sites within CL2 and coil tendencies, power to be at the middle of the α-helix, and refractive index sites within EL2. Using features of the program Swiss-PdbViewer v.3.7 (http://www.expasy.org/spdbv; Guex and Peitsch 1997), selected residues were evaluated for their proximity to the chromophore-binding pocket using bovine rhodopsin as a reference sequence. Using a conservative distance of 4 Å to infer a residue's potential interaction with the chromophore, 10 sites were isolated from those identified by TreeSAAP analyses:

113 (compressibility), 117 (compressibility), 118 (compressibility), 121 (coil tendencies), 122 (compressibility), 186 (power to be at the middle of the α-helix), 187 (power to be at the middle of the α-helix, refractive index), 189 (coil tendencies, power to be at the middle of the α-helix), 207 (coil tendencies, power to be at the middle of the α-helix), and 265 (compressibility). These residues are clustered in 2 areas of the protein in TMIII (113–122) and EL2 (186–207). In addition to these sites, TreeSAAP analyses identified several residues—90 (compressibility, power to

be at the middle of the α-helix), 123 (coil tendencies, compressibility), 164 (compressibility), and 274 (compressibility, power to be at the middle of the α-helix)—which are within ~10 Å of the chromophore and have been identified in other studies of spectral tuning (Wilkie et al. 2000; Briscoe 2002; Salcedo et al. 2003).

Discussion

Pancrustacean Opsin Evolution

Almost all investigated insect opsins fall within 3 spectral clades—long-wavelength, UV, and blue sensitive. These genetic clades fit well with physiological characterizations of spectral sensitivity in insects, supporting the hypothesis that the ancestral visual system was trichromatic (Chittka 1996, 1997). However, although the ancestral state may have been trichromatic, the current diversification of opsin genes appears to be more complex. For example, the complement of 6 visual pigments described from *Drosophila* represent 3 gene duplications that most likely occurred only in the Dipteran lineage (2 duplications leading to the MWS Rh1 and Rh2 and 1 duplicating a UV gene leading to Rh3 and Rh4). More recent studies have identified similar gene complement expansion in insect LWS opsins within specific lineages (Briscoe 1998, 2000; Hill et al. 2002) and early within insect diversification (Spaethe and Briscoe 2004). Although the expression patterns and spectral sensitivities of most of these additional LWS opsin genes have not been investigated, recent studies suggest that at least some of these copies may have an extraocular expression and potentially be involved in circadian regulation systems (Shimizu et al. 2001; Briscoe and White 2005). Similarly, investigations of ostracod opsins have documented differential expression in the median versus the compound eyes as well as the presence of at least 8 expressed SWS loci, indicating many recent gene duplications (Oakley and Huber 2004). Given the demonstrated high copy number of expressed LWS genes in insects and SWS genes in ostracods, a similar pattern of gene expansion is likely to be documented in the crustacean LWS clade once more sequence data are obtained. The low taxonomic clustering observed in the crustacean clade also suggests the presence of unidentified gene duplication events occurring at least before the divergence of the Mysida and Decapoda. Interestingly, however, the crustaceans form 2 main clades that are well supported by BMCMC analyses (p*P* > 0.98). With the exception of *N. oerstedii* Rh1 (489 nm), these 2 clades are roughly divided by spectral sensitivities, with 1 shorter-wavelength clade (496–501 nm) composed of *M. diluviana*, *A. grebnitzkii*, and *E. superba* opsins and 1 longer-wavelength clade (512–533 nm) with sequences from the remaining species. A particularly interesting crustacean group to conduct further investigations of opsin evolution will be the stomatopod crustaceans, here represented by the species *N. oerstedii*, which contain up to 16 different visual pigments that span the ultraviolet to visible spectrum of light (Cronin and Marshall 1989, 2004; Cronin et al. 2000). Characterization of additional LWS opsin sequences will be required to elucidate the validity of the observed patterns of evolution in crustaceans.

An unusual feature of arthropod opsin evolution is the large indel present in the CL3 domain. It exists in all the opsins investigated in this study, with the most conserved stretch taking the form of R(E/D)QAKKM(N/G). Because the function of the CL3 domain has been linked to G protein docking and signal transfer, and a relatively large proportion of its sites were identified to be under destabilizing positive selection for compressibility (16.1%) and refractive index (22.2%) (fig. 4), this conserved region should be of interest to future comparisons of opsin functionality between "invertebrates" and vertebrates.

Selective Forces in Pancrustacean Opsins

The ability to study the genetic mechanisms behind phenotypic variation has made opsin a model evolutionary system for studying the selective influences leading to genetic adaptation. The combination of genetic, physiologic, and biochemical studies have shown a direct correlation between the environment and visual pigment spectral sensitivity (Lythgoe 1972, 1980; Crescitelli et al. 1985; Partridge et al. 1988, 1989; Partridge 1989; Douglas et al. 1998) and demonstrated the effects of single amino acid changes on tuning this spectral sensitivity (Nathans 1990a; Neitz et al. 1991; Asenjo et al. 1994; Yokoyama and Radlwimmer 1999; Wilkie et al. 2000; Yokoyama et al. 2000; Yokoyama and Radlwimmer 2001; Cowing et al. 2002; Yokoyama and Tada 2003), making opsins one of the clearest, and best-studied, protein groups for the quantification of adaptation at the molecular level. Because the observed variation in visual pigment spectral absorbance is the result of adaptive selection on the opsin gene, this gene family is therefore an excellent system for testing methods of detecting how selection acts on protein sequences. In our analyses, estimated d_N/d_S ratios did not detect any evidence of selection; however, these results only confirm that the opsin gene is a generally conservative protein-coding gene that requires the use of alternative criteria to reveal molecular adaptations (McClellan et al. 2005). Using a method that evaluates only nonsynonymous replacements with respect to the relative change in a suite of amino acid properties, 4 properties were identified to be under positive destabilizing selection across the phylogeny: coil tendencies, compressibility, power to be at the middle of the α-helix, and refractive index. Interestingly, none of the identified properties measure aspects of polarity or charge, 2 of the properties hypothesized to be important in residues affecting vertebrate spectral tuning (Nathans 1987, 1990b; Neitz et al. 1991; Chan et al. 1992). Instead, the identified properties mostly relate to structural aspects. Thus, although polarity and charge may be important in fine-tuning spectral absorbance, in a broad-scale evolutionary context, structural aspects of residue replacement seems to have been more important to overall opsin function. In particular, compressibility seems to be a significant property impacting opsin evolution, with 186 nonsynonymous replacements identified from across the phylogeny, with all replacements being to the residue alanine. Alanine is 1 of 7 amino acids identified by Gromiha and Ponnuswamy (1993) that determine overall protein compressibility, and, with the exception of proline, has the smallest calculated

compressibility (-25.5×10^{-15} m^3 mol^{-1} Pa^{-1}). Previous studies have indicated that alanine replacements can stabilize α-helices and have implicated alanine content in helical thermal stability (Argos et al. 1979; Lyu et al. 1990; O'Neil and DeGrado 1990; Padmanabhan et al. 1990; Zhang et al. 1991; Ptitsyn 1992; although see Pinker et al. 1993 for a different opinion). The large proportion of nonsynonymous replacements in TMI (10.2%), TMIII (17.7%), TMIV (18.3%), TMV (3.8%), and TMVI (13.4%) indicates that α-helical stability may play an important role in pancrustacean opsin evolution. The exclusive use of alanine over proline at these sites is most likely due to the fact that proline residues introduce bends into α-helices (Riek et al. 2001).

Most amino acid sites corresponding to the 4 identified properties occur in transmembrane domains (fig. 4). Based on site-directed mutagenesis studies in vertebrates, transmembrane domains III, VI, and VII are known to have significant interaction with the chromophore, and interactions among these 3 helices are involved in restraining the structure of GPCRs in the inactive, nonsignaling state (Sakmar et al. 1989; Zhukovsky and Oprian 1989; Nathans 1990a, 1990b; Nakayama and Khorana 1991; Chan et al. 1992; Filipek et al. 2003). Additionally, the cytoplasmic ends of helices III, VI, and VII move during GPCR activation to create a binding crevice for the G protein (Filipek et al. 2003). Unfortunately, the selective influences in TMVII could not be evaluated in this study because many of the available opsin sequences are too short. However, the fact that 3 of the properties (coil tendencies, compressibility, and power to be at the middle of the α-helix) had significant amounts of detected selection in TMIII indicates that this helix may play a crucial role in the functional diversification of arthropod opsins.

Using a cutoff distance of 4 Å, 10 sites from the TreeSAAP analyses were identified that potentially affect the chromophore-binding pocket. This criterion is very conservative given that residues at distances of up to 10 Å have been identified as regulating chromophore wavelength absorption in other studies of spectral tuning (Wilkie et al. 2000; Briscoe 2002; Salcedo et al. 2003). In the bovine rhodopsin structure, 18 residues are within 4 Å of the chromophore; we detected destabilizing selection at 55.6% of those residues. Comparisons of these sites with studies using evolutionary trace (ET) analyses to identify transmembrane residues of importance to class A GPCRs in general and to opsins specifically indicate that at least 2 of our identified residues (113 and 207) correspond to previously identified residues specific to opsin function, whereas an additional 3 residues (118, 121, and 265) are of general importance to GPCR functioning (Madabushi et al. 2004). Our analyses identified an additional 2 transmembrane residues (117 and 122) to be under destabilizing selection, which were not found using the ET methods, as well as 3 residues located in EL2, which were not included in the ET analyses. This comparison illustrates that TreeSAAP methods are capable of detecting selection at residues of importance to general GPCR- and opsin-specific functioning, while also finding novel sites not predicted using other methods. Further Tree-SAAP analyses–comparing sites found in opsins versus broader sets of class A GPCRs are needed to tease apart the residues that are most important to opsins.

The 10 residues identified to be under destabilizing selection and within 4 Å of the chromophore are clustered in the extracellular end of TMIII and in EL2. The EL2 folds back into the cavity formed by the membrane-embedded domains, forming part of the chromophore-binding pocket and acting as a "plug" preventing solvent access to the Schiff base (Yan et al. 2003). Recent studies have shown this loop is important for the thermal stability of the dark state of rhodopsin (Janz et al. 2003), and in particular, E181 in bovine rhodopsin affects the stability and spectral absorption of metarhodopsin II (Yan et al. 2003), leading to the hypothesis that E181 is a counterion for metarhodopsin I (Teller et al. 2003; Yan et al. 2003). Interestingly, although the rhodopsin counterion E113 is not conserved outside of vertebrates, the E181 residue is conserved, implying that there may be more similarities in metarhodopsin dynamics between nonvertebrates and vertebrates than in rhodopsin spectral tuning. All of the sites identified in this region (186, 187, and 189) demonstrated positive destabilizing selection for power to be at the middle of the α-helix, suggesting that similar forces are acting upon adaptation of residues in the EL2 "plug."

The sites clustered at the extracellular end of TMIII (113, 117, 118, 121, and 122) were demonstrated to be under selection for compressibility. These sites are all on turns of the α-helix that face the chromophore-binding pocket, and several of these sites have been identified as affecting spectral tuning in vertebrates (Nathans 1990a, 1990b; Yokoyama et al. 1999; Yokoyama and Tada 2000; Shi et al. 2001; Yokoyama and Tada 2003). Intriguingly, site 113 (the vertebrate counterion site) was identified to be under positive destabilizing selection in this data set. This site is occupied by either a tyrosine in visible-absorbing or a phenylalanine in UV-absorbing arthropod pigments. However, previous studies have shown this site is not used as a counterion (Nakagawa et al. 1999) and that the observed amino acid polymorphism is not responsible for the difference in absorption spectra between UV and visible pigments (Salcedo et al. 2003). However, the identification of the site in this study implies that the observed polymorphism serves unidentified function in visible versus UV opsins.

In addition to the sites in TMIII and EL2, several residues (90, 123, and 274) identified in our analyses to be under positive destabilizing selection for more than 1 property are within ~10 Å of the chromophore and have been identified in other studies of spectral tuning (Wilkie et al. 2000; Briscoe 2001, 2002; Salcedo et al. 2003). The identification of these sites confirms that amino acids affecting spectral tuning are selected for impacts on structural aspects of the helices (in this study, compressibility, power to be at the middle of the α-helix, and coil tendencies). These amino acid properties in particular may affect the internal packing of the chromophore-binding site and thereby affect spectral tuning of the chromophore and signal transduction by the opsin protein.

Summary

In this study, we contribute to the knowledge of "pancrustacean" opsin evolution by identifying amino acid sites within the chromophore-binding pocket under selection.

Techniques investigating amino acid property changes detected 4 amino acid properties (coil tendencies, compressibility, power to be at the middle of an α-helix, and refractive index) to be historically influenced by destabilizing positive selection. Ten amino acid sites relating to these properties were found to be facing the binding pocket, within 4 Å of the chromophore, with potential to affect spectral tuning. Comparisons with previous studies investigating opsin functional residues indicate that some of these sites are specific to opsin functioning, some are general to GPCR functioning, and some are novel residues not previously identified. This research demonstrates the ability of TreeSAAP to identify amino acid residues of functional importance in an evolutionary context, providing a method for targeting sites for further studies, particularly in proteins that have not yet been structurally determined.

Acknowledgments

Literature Cited

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics. 21:2104–2105.

Anisimova M, Bielawski JP, Yang Z. 2002. Accuracy and power of Bayes prediction of amino acid sites under positive selection. Mol Biol Evol. 19:950–958.

Arendt D, Tessmar-Raible K, Snyman H, Dorresteijn AW, Wittbrodt J. 2004. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. Science. 306:869–871.

Argos P, Rossmann MG, Grau UM, Zuber H, Frank G, Tratschin JD. 1979. Thermal stability and protein structure. Biochemistry. 18:5698–5703.

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

Arikawa K, Scholten DGW, Kinoshita M, Stavenga DG. 1999. Tuning of photoreceptor spectral sensitivities by red and yellow pigments in the butterfly *Papilio xuthus*. Zool Sci. 16:17–24.

Asenjo AB, Rim J, Oprian DD. 1994. Molecular determinants of human red/green color discrimination. Neuron. 12:1131–1138.

Bellingham J, Wells DJ, Foster RG. 2003. In silico characterisation and chromosomal localisation of human *Rrh* (peropsin)—implications for opsin evolution. BMC Genomics. 4:3.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The protein data bank. Nucleic Acids Res. 28:235–242.

Bielawski JP, Yang Z. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. J Mol Evol. 59:121–132.

Bracho MA, Moya A, Barrio E. 1998. Contributions of *Taq* polymerase-induced errors to the estimation of RNA virus diversity. J Gen Virol. 79:2921–2928.

Briscoe AD. 1998. Molecular diversity of visual pigments in the butterfly *Papilio glaucus*. Naturwissenschaften. 85:33–35.

Briscoe AD. 2000. Six opsins from the butterfly *Papilio glaucus*: molecular phylogenetic evidence for paralogous origins of red-sensitive visual pigments in insects. J Mol Evol. 51:110–121.

Briscoe AD. 2001. Functional diversification of lepidopteran opsins following gene duplication. Mol Biol Evol. 18:2270–2279.

Briscoe AD. 2002. Homology modeling suggests a functional role for parallel amino acid substitutions between bee and butterfly red- and green-sensitive opsins. Mol Biol Evol. 19:983–986.

Briscoe AD, Bernard GD, Szeto AS, Nagy LM, White RH. 2003. Not all butterfly eyes are created equal: rhodopsin absorption spectra, molecular identification, and localization of ultraviolet-, blue-, and green-sensitive rhodopsin-encoding mRNAs in the retina of *Vanessa cardui*. J Comp Neurol. 458:334–349.

Briscoe AD, Chittka L. 2001. The evolution of color vision in insects. Annu Rev Entomol. 46:471–510.

Briscoe AD, White RH. 2005. Adult stemmata of the butterfly *Vanessa cardui* express UV and green opsin mRNAs. Cell Tissue Res. 319:175–179.

Britt SG, Feiler R, Kirschfeld K, Zuker CS. 1993. Spectral tuning of rhodopsin and metarhodopsin in vivo. Neuron. 11:29–39.

Brown AJH. 1996. Isolation and characterisation of visual pigment genes from the stomatopod crustacean *Gonodactylus oerstedii* [Ph.D. dissertation]. Sussex, UK: University of Sussex.

Brown PK, Brown PS. 1958. Visual pigments of the octopus and cuttlefish. Nature. 182:1288–1290.

Carulli JP, Chen DM, Stark WS, Hartl DL. 1994. Phylogeny and physiology of *Drosophila* opsins. J Mol Evol. 38:250–262.

Carulli JP, Hartl DL. 1992. Variable rates of evolution among *Drosophila* opsin genes. Genetics. 132:193–204.

Chan T, Lee M, Sakmar TP. 1992. Introduction of hydroxyl-bearing residues causes bathochromic spectral shifts in rhodopsin. J Biol Chem. 267:9478–9480.

Chang BSW, Ayers D, Smith WC, Pierce NE. 1996. Cloning of the gene encoding honeybee long-wavelength rhodopsin: a new class of insect visual pigments. Gene. 173:215–219.

Chang BSW, Crandall KA, Carulli JP, Hartl DL. 1995. Opsin phylogeny and evolution: a model for blue shifts in wavelength regulation. Mol Phylogenet Evol. 4:31–43.

Chase MR, Bennett RR, White RH. 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.

Chittka L. 1997. Bee color vision is optimal for coding flower color, but flower colors are not optimal for being coded—why? Israel J Plant Sci. 45:115–127.

Cowing JA, Poopalasundaram S, Wilkie SE, Robinson PR, Bowmaker JK, Hunt DM. 2002. The molecular mechanism for the spectral shifts between vertebrate ultraviolet- and violet-sensitive cone visual pigments. Biochem J. 367:129–135.

Crandall KA, Cronin TW. 1997. The molecular evolution of visual pigments of freshwater crayfishes (Decapoda: Cambaridae). J Mol Evol. 45:524–534.

Crandall KA, Hillis DM. 1997. Rhodopsin evolution in the dark. Nature. 387:667–668.

Crescitelli F, McFall-Ngai M, Horwitz J. 1985. The visual pigment sensitivity hypothesis: further evidence from fishes of varying habitats. J Comp Physiol A. 157:323–333.

Cronin TW, Goldsmith TH. 1982. Photosensitivity spectrum of crayfish rhodopsin measured using fluorescence of metarhodopsin. J Gen Physiol. 79:313–332.

Cronin TW, Marshall NJ. 1989. Multiple spectral classes of photoreceptors in the retinas of gonodactyloid stomatopod crustaceans. J Comp Physiol A. 166:261–275.

Cronin TW, Marshall NJ. 2004. The unique visual world of mantis shrimps. In: Prete F, editor. Complex worlds from simple nervous systems. Cambridge (MA): MIT Press. p. 239–268.

Cronin TW, Marshall NJ, Caldwell RL. 1994. The retinas of mantis shrimps from low-light environments (Crustacea, Stomatopoda, Gonodactylidae). J Comp Physiol A. 174:607–619.

Cronin TW, Marshall NJ, Caldwell RL. 1996. Visual pigment diversity in two genera of mantis shrimps implies rapid evolution (Crustacea; Stomatopoda). J Comp Physiol A. 179:371–384.

Cronin TW, Marshall NJ, Caldwell RL. 2000. Spectral tuning and the visual ecology of mantis shrimps. Phil Trans R Soc Lond B. 355:1263–1267.

Douglas RH, Partridge JC, Marshall NJ. 1998. The eyes of deep-sea fish I. Lens pigmentation, tapeta and visual pigments. Prog Retin Eye Res. 17:597–636.

Eguchi E, Watanabe K, Hariyama T, Yamamoto K. 1982. A comparison of electrophysiologically determined spectral responses in 35 species of Lepidoptera. J Insect Physiol. 28:675–682.

Feiler R, Bjornson R, Kirschfeld K, Mismer D, Rubin GM, Smith DP, Socolich M, Zuker CS. 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 WA, Kirschfeld K, Wehrhahn C, Zuker CS. 1988. Targeted misexpression of a *Drosophila* opsin gene leads to altered visual function. Nature. 333:737–741.

Filipek S, Teller DC, Palczewski K, Stenkamp R. 2003. The crystallographic model of rhodopsin and its use in studies of other G protein-coupled receptors. Annu Rev Biophys Biomol. 32:375–397.

Forsberg R, Christiansen FB. 2003. A codon-based model of host-specific selection in parasites, with an application to the influenza A virus. Mol Biol Evol. 20:1252–1259.

Frank TM, Widder EA. 1999. Comparative study of the spectral sensitivities of mesopelagic crustaceans. J Comp Physiol A. 185:255–265.

Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol. 63:1256–1272.

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

Goldman LJ, Barnes SN, Goldsmith TH. 1975. Microspectrophotometry of rhodopsin and metarhodopsin in the moth *Galleria*. J Gen Physiol. 66:383–404.

Goldsmith TH. 1978. The spectral absorption of crayfish rhabdoms: pigment, photoproduct and pH sensitivity. Vision Res. 18:463–473.

Gromiha MM, Ponnuswamy PK. 1993. Relationship between amino-acid properties and protein compressibility. J Theor Biol. 165:87–100.

Guex N, Peitsch MC. 1997. Swiss-model and the swiss-pdb viewer: an environment for comparative protein modeling. Electrophoresis. 18:2714–2723.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52:696–704.

Hall MD, Hoon MA, Ryba NJ, Pottinger JD, Keen JN, Saibil HR, Findlay JB. 1991. Molecular cloning and primary structure of squid (*Loligo forbesi*) rhodopsin, a phospholipase C-directed G-protein linked receptor. Biochem J. 274:35–40.

Hill CA, Fox AN, Pitts RJ, Kent LB, Tan PL, Chrystal MA, Cravchik A, Collins FH, Robertson HM, Zwiebel LJ. 2002. G protein-coupled receptors in *Anopheles gambiae*. Science. 298:176–178.

Hubbard R, St George RCC. 1958. The rhodopsin system of the squid. J Gen Physiol. 41:510–528.

Hubbard R, Wald G. 1960. Visual pigment of the horseshoe crab, *Limulus polyphemus*. Nature. 186:212–215.

Huelsenbeck JP, Imennov NS. 2002. Geographic origin of human mitochondrial DNA: accomodating phylogenetic uncertainty and model comparison. Syst Biol. 51:155–165.

Huelsenbeck JP, Larget B, Miller RE, Ronquist F. 2002. Potential applications and pitfalls of Bayesian inference phylogeny. Syst Biol. 51:673–688.

Ichikawa T, Tateda H. 1982. Distribution of color receptors in the larval eyes of four species of Lepidoptera. J Comp Physiol A. 149:317–324.

Janz JM, Fay JF, Farrens DL. 2003. Stability of dark state rhodopsin is mediated by a conserved ion pair in intradiscal loop E-2. J Biol Chem. 278:16982–16991.

Jokela-Määttä M, Pahlberg J, Lindström M, Zak PP, Porter ML, Ostrovskii MA, Cronin TW, Donner K. 2005. Visual pigment absorbance and spectral sensitivity of the *Mysis relicta* species group (Crustacea, Mysidae) in different light environments. J Comp Physiol A. 191:1087–1097.

Kent J. 1997. The visual pigments of deep-sea crustaceans [Ph.D. dissertation]. Sussex (UK): University of Bristol.

Kitamoto JK, Sakamoto K, Ozaki K, Mishina Y, Arikawa K. 1998. Two visual pigments in a single photoreceptor cell: identification and histological localization of three mRNAs encoding visual pigment opsins in the retina of the butterfly *Papilio xuthus*. J Exp Biol. 201:1255–1261.

Koutalos Y, Ebrey TG, Tsuda M, et al. (11 co-authors). 1989. Regeneration of bovine and octopus opsins in situ with natural and artificial retinals. Biochemistry. 28:2732–2739.

Langer H, Hamann B, Meinecke CC. 1979. Tetrachromatic visual system in the moth *Spodoptera exempta* (Insecta: Noctuidae). J Comp Physiol A. 129:235–239.

Lythgoe JN. 1972. The adaptation of visual pigments to the photic environment. In: Dartnall HJA, editor. Handbook of sensory physiology. Berlin (Germany): Springer-Verlag. p. 566–603.

Lythgoe JN. 1980. Vision in fish: ecological adaptations. In: Ali MA, editor. Environmental physiology of fishes. New York: Plenum Press. p. 431–445.

Lyu PC, Liff MI, Marky LA, Kallenback NR. 1990. Side chain contributions to the stability of alpha helical structure in peptides. Science. 250:669–673.

Madabushi S, Gross AK, Philippi A, Meng EC, Wensel TG, Lichtarge O. 2004. Evolutionary trace of G protein-coupled receptors reveals clusters of residues that determine global and class-specific functions. J Biol Chem. 279:8126–8132.

Maddison DR, Maddison WP. 2001. MacClade 4. Analysis of phylogeny and character evolution. Version 4.05. Sunderland (MA): Sinauer Associates.

McClellan DA, McCracken KG. 2001. Estimating the influence of selection on the variable amino acid sites of the cytochrome b protein functional domains. Mol Biol Evol. 18:917–925.

McClellan DA, Palfreyman EJ, Smith MJ, Moss JL, Christensen RG, Sailsbery JK. 2005. Physiocochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome *b* proteins. Mol Biol Evol. 22:437–455.

Montell C, Jones K, Zuker CS, Rubin GM. 1987. A second opsin gene expressed in the ultraviolet-sensitive R7 photoreceptor cells of *Drosophila melanogaster*. J Neurosci. 7:1558–1566.

Morris A, Bowmaker JK, Hunt DM. 1993. The molecular basis of a spectral shift in the rhodopsins of two species of squid from different photic environments. Proc R Soc Lond B. 254:233–240.

Naito T, Nashima-Hayama K, Ohtsu K, Kito Y. 1981. Photoreactions of cephalopod rhodopsin. Vision Res. 21:935–941.

Nakagawa M, Iwasa T, Kikkawa S, Tsuda M, Ebrey TG. 1999. How vertebrate and invertebrate visual pigments differ in their mechanism of photoactivation. Proc Natl Acad Sci USA. 96:6189–6192.

Nakayama TA, Khorana HG. 1991. Mapping of the amino acids in membrane-embedded helices that interact with the retinal chromophore in bovine rhodopsin. J Biol Chem. 266:4269–4275.

Nathans J. 1987. Molecular biology of visual pigments. Ann Rev Neurosci. 10:163–194.

Nathans J. 1990a. Determinants of visual pigment absorbance: identification of the retinylidene 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 putative transmembrane segments. Biochemistry. 29:937–942.

Neitz M, Neitz J, Jacobs GH. 1991. Spectral tuning of pigments underlying red-green color vision. Science. 252:971–974.

Nielsen R, Yang ZH. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics. 148:929–936.

Nolte J, Brown JE. 1972. Electrophysiological properties of cells in the median ocellus of *Limulus*. J Gen Physiol. 59:167–185.

Oakley TH, Huber DR. 2004. Differential expression of duplicated opsin genes in two eye types of ostracod crustaceans. J Mol Evol. 59:239–249.

O'Neil KT, DeGrado WF. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. Science. 250:646–651.

O'Tousa JE, Baehr W, Martin RL, Hirsh J, Pak WL, Applebury ML. 1985. The *Drosophila ninaE* gene encodes an opsin. Cell. 40:839–850.

Padmanabhan S, Marqusee S, Ridgeway T, Laue TM, Baldwin RL. 1990. Relative helix-forming tendencies of nonpolar amino acids. Nature. 344:268–270.

Palczewski K, Kumasaka T, Hori T, et al. (12 co-authors). 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. Science. 289:739–745.

Partridge JC. 1989. Visual pigments of deep-sea fishes: ecophysiology and molecular biology. Prog Underw Sci. 14:17–31.

Partridge JC, Archer SN, Lythgoe JN. 1988. Visual pigments in the individual rods of deep-sea fishes. J Comp Physiol A. 162:543–550.

Partridge JC, Shand J, Archer SN, Lythgoe JN, van Groningen-Luyben WA. 1989. Interspecific variation in the visual pigments of deep-sea fishes. J Comp Physiol A. 164:513–529.

Paul R, Steiner A, Gemperlein R. 1986. Spectral sensitivity of *Calliphora erythrocephala* and other insect species studied with fourier interferometric stimulation (fis). J Comp Physiol A. 158:669–680.

Peitsch D, Fietz A, Hertel H, de Souza J, Ventura DF, Menzel R. 1992. The spectral input systems of hymenopteran insects and their receptor-based colour vision. J Comp Physiol A. 170:23–40.

Pichaud F, Briscoe A, Desplan C. 1999. Evolution of color vision. Curr Opin Neurobiol. 9:622–627.

Pinker RJ, Lin L, Rose GD, Kallenbach NR. 1993. Effects of alanine substitution in alpha-helices of sperm whale myoglobin on protein stability. Protein Sci. 2:1099–1105.

Popp MP, Grisshammer R, Hargrave PA, Smith WC. 1996. Ant opsins: sequences from the saharan silver ant and the carpenter ant. Invertebr Neurosci. 1:323–329.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst Biol. 53:793–808.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics. 14:817–818.

Ptitsyn OB. 1992. Secondary structure formation and stability. Curr Opin Struc Biol. 2:13–20.

Rambaut A, Drummond AJ. 2003. Tracer. Version 1.2 [Internet]. Oxford (UK): University of Oxford [cited 2005 May 31]. Available from: http://evolve.zoo.ox.ac.uk.

Riek RP, Rigoutsos I, Novotny J, Graham RM. 2001. Non-alpha-helical elements modulate polytopic membrane protein architecture. J Mol Evol. 306:349–362.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:1572–1574.

Rossel S. 1979. Regional differences in photoreceptor performance in the eye of the praying mantis (*Tenodera australasiae*). J Comp Physiol A. 131:95–112.

Saiki R, Gelfand DH, Stofell S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 239:487–491.

Sakamoto K, Hisatomi O, Tokunaga F, Eguchi E. 1996. Two opsins from the compound eye of the crab *Hemigrapsus sanguineus*. J Exp Biol. 199:441–450.

Sakmar TP, Franke RR, Khorana HG. 1989. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. Proc Natl Acad Sci USA. 86:8309–8313.

Salcedo E, Huber A, Henrich S, Chadwell LV, Chou WH, Paulsen R, Britt SG. 1999. Blue- and green-absorbing visual pigments of *Drosophila*: ectopic expression and physiological characterization of the R8 photoreceptor cell-specific Rh5 and Rh6 rhodopsins. J Neurosci. 19:10716–10726.

Salcedo E, Zheng L, Phistry M, Bagg EE, Britt SG. 2003. Molecular basis for ultraviolet vision in invertebrates. J Neurosci. 23:10873–10878.

Shi Y, Radlwimmer FB, Yokoyama S. 2001. Molecular genetics and the evolution of ultraviolet vision in vertebrates. Proc Natl Acad Sci USA. 98:11731–11736.

Shimizu I, Yamakawa Y, Shimazaki Y, Iwasa T. 2001. Molecular cloning of *Bombyx* cerebral opsin (boceropsin) and cellular localization of its expression in the silkworm brain. Biochem Biophys Res Commun. 287:27–34.

Smith WC, Ayers DM, Popp MP, Hargrave PA. 1997. Short wavelength-sensitive opsins from the Saharan silver and carpenter ants. Invert Neurosci. 3:49–56.

Smith WC, Price DA, Greenberg RM, Battelle BA. 1993. Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus polyphemus*. Proc Natl Acad Sci USA. 90:6150–6154.

Spaethe J, Briscoe AD. 2004. Early duplication and functional diversification of the opsin gene family in insects. Mol Biol Evol. 21:1583–1594.

Stavenga DG, Smits RP, Hoenders BJ. 1993. Simple exponential functions describing the absorbance bands of visual pigment spectra. Vision Res. 33:1011–1017.

Struwe G. 1972. Spectral sensitivity of the compound eye in butterflies (*Heliconius*). J Comp Physiol A. 79:191–196.

Taylor SD, Dittmar de la Cruz K, Porter ML, Whiting MF. 2005. Characterization of the long-wavelength opsin from Mecoptera and Siphonaptera: does a flea see? Mol Biol Evol. 22:1165–1174.

Teller DC, Stenkamp RE, Palczewski K. 2003. Evolutionary analysis of rhodopsin and cone pigments: connecting the three-dimensional structure with spectral tuning and signal transfer. FEBS Lett. 555:151–159.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.

Towner P, Gartner W. 1994. The primary structure of mantid opsin. Genetics. 143:227–231.

Townson SM, Chang BSW, Salcedo E, Chadwell LV, Pierce NE, Britt SG. 1998. Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in *Drosophila*, and physiological characterization. J Neurosci. 18:2412–2422.

Vanhoutte KJA, Eggen BJL, Janssen JJM, Stavenga DG. 2002. Opsin cDNA sequences of a UV and green rhodopsin of the satyrine butterfly *Bicyclus anynana*. Insect Biochem Mol Biol. 32:1383–1390.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol. 18:691–699.

White RH, Brown PK, Hurley AK, Bennett RR. 1983. Rhodopsins, retinula cell ultrastructure, and receptor potentials in the developing pupal eye of the moth *Manduca sexta*. J Comp Physiol A. 150:153–163.

Wilkie SE, Robinson PR, Cronin TW, Poopalasundaram S, Bowmaker JK, Hunt DM. 2000. Spectral tuning of avian violet- and ultraviolet-sensitive visual pigments. Biochemistry. 39:7895–7901.

Woolley S, Johnson J, Smith MJ, Crandall KA, McClellan DA. 2003. TreeSAAP: selection on amino acid properties using phylogenetic trees. Bioinformatics. 19:671–672.

Yan EC, Kazmi MA, De S, Chang BSW, Seibert C, Marin EP, Mathies RA, Sakmar TP. 2003. Function of extracellular loop 2 in rhodopsin: glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II. Biochemistry. 41:3620–3627.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13:555–556.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol. 15:568–573.

Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J Mol Evol. 46:409–418.

Yang Z, Nielsen R, Goldman N, Pedersen AMK. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 155:431–449.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol. 22:1107–1118.

Yokoyama S, Radlwimmer FB. 1999. The molecular genetics of red and green color vision in mammals. Genetics. 153:919–932.

Yokoyama S, Radlwimmer FB. 2001. The molecular genetics and evolution of red and green color vision in vertebrates. Genetics. 158:1697–1710.

Yokoyama S, Radlwimmer FB, Blow NS. 2000. Ultraviolet pigments in birds evolved from violet pigments by a single amino acid change. Proc Natl Acad Sci USA. 97:7366–7371.

Yokoyama S, Tada T. 2000. Adaptive evolution of the African and Indonesian coelacanths to deep-sea environments. Gene. 261:35–42.

Yokoyama S, Tada T. 2003. The spectral tuning in the short wavelength-sensitive type 2 pigments. Gene. 306:91–98.

Yokoyama S, Zhang H, Radlwimmer FB, Blow NS. 1999. Adaptive evolution of color vision of the Comoran coelacanth (*Latimeria chalumnae*). Proc Natl Acad Sci USA. 96:6279–6284.

Zeiger J, Goldsmith TH. 1994. Behavior of crayfish rhodopsin and metarhodopsin in digitonin: the 510 and 562 nm 'visual pigments' are artifacts. Vision Res. 34:2679–2688.

Zhang XJ, Baase WA, Matthews BW. 1991. Toward a simplification of the protein folding problem: a stabilizing polyalanine alpha-helix engineered in T4 lysozyme. Biochemistry. 30:2012–2017.

Zhukovsky EA, Oprian DD. 1989. Effect of carboxylic acid side chains on the absorption maximum of visual pigments. Science. 246:928–930.

Zhukovsky EA, Robinson PR, Oprian DD. 1992. Changing the location of the Schiff base counterion in rhodopsin. Biochemistry. 31:10400–10405.

Zuker CS, Montell C, Jones K, Laverty T, Rubin GM. 1987. A rhodopsin gene expressed in photoreceptor cell R7 of the *Drosophila* eye: homologies with other signal-transducing molecules. J Neurosci. 7:1550–1557.