The Journal of Experimental Biology 200, 2469–2478 (1997)
Printed in Great Britain © The Company of Biologists Limited 1997
JEB1083

THREE OPSIN-ENCODING cDNAS FROM THE COMPOUND EYE OF *MANDUCA SEXTA*

MICHAEL R. CHASE, RUTH R. BENNETT AND RICHARD H. WHITE*
Department of Biology, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA 02125-3393, USA

Accepted 10 July 1997

Summary

Three distinct opsin-encoding cDNAs, designated MANOP1, MANOP2 and MANOP3, were isolated from the retina of the sphingid moth *Manduca sexta*. MANOP1 codes for a protein with 377 amino acid residues. It is similar in sequence to members of a phylogenetic group of long-wavelength-sensitive arthropod photopigments, most closely resembling the opsins of ants, a praying mantis, a locust and the honeybee. MANOP2 and MANOP3 opsins have 377 and 384 residues respectively. They belong to a related group of insect visual pigments that include the ultraviolet-sensitive rhodopsins of flies as well as other insect rhodopsins that are also thought to absorb at short wavelengths. The retina of *Manduca sexta* contains three rhodopsins, P520, P450 and P357, with absorbance peaks, respectively, at green, blue and ultraviolet wavelengths. There is evidence that MANOP1 encodes the opsin of P520. We suggest that MANOP2 encodes P357 and that MANOP3, representing a class of blue-sensitive insect photopigments, encodes P450.

Key words: Lepidoptera, Sphingidae, *Manduca sexta*, compound eye, rhodopsin, opsin cDNA.

Introduction

Visual pigments (generically, rhodopsins) are intrinsic membrane proteins, opsins, coupled with retinoid chromophores. They are functionally distinguished by the wavelength of maximal absorbance (λ_{max}). There is good reason to believe that amino acid residues in the opsin-binding pocket surrounding the chromophore interact with its unsaturated polyene chain to determine the λ_{max} of that particular rhodopsin. However, the molecular details are far from settled (Yokoyama, 1995; Yokoyama and Yokoyama, 1996).

The presence in an eye of two or more rhodopsins tuned to different regions of the spectrum can enable an animal to discriminate relevant visual stimuli on the basis of wavelength. Behavioral, electrophysiological and biochemical studies indicate that wavelength discrimination in insects is commonly based on a set of three rhodopsins tuned to absorb maximally in the ultraviolet, blue and green regions of the spectrum (Menzel and Backhaus, 1991; Peitsch *et al.* 1992). At least in Hymenoptera, this capacity fits the definition of true trichromatic color vision (Goldsmith, 1990; Menzel and Backhaus, 1991; Peitsch *et al.* 1992). Some insects also have additional red receptors (Bernard and Remington, 1991; Peitsch *et al.* 1992).

Vertebrate, cephalopod and arthropod opsins occupy distinct branches of a phylogenetic tree that may have separated as early as the Paleozoic period (Goldsmith, 1990; Gärtner and Towner, 1995). The sequences of nearly 50 assorted vertebrate opsins are now available (Yokoyama and Yokoyama, 1996). In contrast, only a few insect (and other invertebrate) opsins are known (for references and sequence comparisons, see Smith *et al.* 1993; Carulli *et al.* 1994; Chang *et al.* 1995; Gärtner and Towner, 1995). The scant comparative evidence suggests that the molecular basis of spectral tuning (apart from very general considerations) may be dissimilar in insect and vertebrate rhodopsins (Britt *et al.* 1993; Chang *et al.* 1995; Gärtner and Towner, 1995). Thus, opsin amino acid sequences provide the primary data for attacking several related problems: the molecular basis of spectral tuning, the phylogeny of visual pigments and the evolution of color vision. The primary structures of insect visual pigments promise, in addition, a valuable comparative perspective provided by the molecular products of parallel evolution.

It is not surprising that most of our information on insect opsins comes from flies. Four opsin-encoding genes from *Drosophila melanogaster* have been known for some time (Pollock and Benzer, 1988). Each is expressed in specific photoreceptor cells of the compound eye, the dorsal ocelli and the larval photosensory organ. The opsins they encode fall into two groups on the basis of amino acid sequence and rhodopsin

*Author for correspondence (e-mail: rwhite@umbsky.cc.umb.edu).

absorbance maxima: Rh1 and Rh2 encode rhodopsins that absorb in the blue-green region of the visible spectrum. Rh3 and Rh4 encode ultraviolet-sensitive rhodopsins. Two more distinctive *Drosophila melanogaster* opsin sequences, Rh5 (Chou *et al.* 1996) and Rh6 (Huber *et al.* 1997), have just been discovered. Efforts to characterize opsins from insect taxa other than Diptera have yielded sequences from a mantid species (Towner and Gärtner, 1994), the locust *Schistocerca gregaria* (Gärtner and Towner, 1995; Towner *et al.* 1997), two species of ant, *Cataglyphis bombycina* and *Camponotus abdominalis* (Popp *et al.* 1996), and the honeybee *Apis mellifera* (Chang, 1995; Chang *et al.* 1996; Bellingham *et al.* 1997). However, the corresponding rhodopsins have not been directly identified from the absorbance spectra of expressed gene products or by *in situ* localization.

To explore systematically the lineage and function of insect visual pigments requires the continued collection of opsin sequences from insect orders other than Diptera. It is particularly desirable to characterize the full complement of opsins from compound eyes that deploy the three photopigments of typical λ_{max} for wavelength discrimination. The sphingid moth *Manduca sexta* has such a representative trichromatic visual system based on characterized rhodopsins, P520, P450 and P357, absorbing in the green, blue and ultraviolet regions respectively (White *et al.* 1983*a*,*b*; Bennett and Brown, 1985) and it supports wavelength-dependent behavior (White *et al.* 1994; Cutler *et al.* 1995). The present paper reports the amino acid sequences of three opsins deduced from retinal cDNA and addresses the problem of assigning them among the three identified rhodopsins.

Materials and methods

Isolation and cloning of opsin-encoding cDNA

Manduca sexta were reared on a carotenoid-rich artificial diet under conditions described previously (Bennett and White, 1989; Chase *et al.* 1996). Approximately 200 retinas from compound eyes were frozen quickly in liquid nitrogen and stored at −80 °C. Total RNA was isolated using guanidinium thiocyanate extraction and pelleted twice by successive centrifugation through 5.7 mol l^{-1} CsCl (MacDonald *et al.* 1987) in order to remove traces of ommachrome screening pigment and/or other factors that may interfere with reverse transcription (Smith *et al.* 1993).

Total RNA was reverse-transcribed using a primer (RT-1, Table 1) targeted to poly(A) in order to select mRNA. Approximately 200 ng of RNA and 100 pmol of the primer were added to 50 µl of 5× reverse transcription buffer (Gibco BRL, Rockville, Maryland, USA). The solution was heated to 65 °C and allowed to cool to room temperature for 10 min. Reverse Transcriptase RNase Minus (100 units; M-MLV, Gibco) and 1 mmol l^{-1} each of dATP, dCTP, dGTP and dTTP were then added, and the reverse transcription reaction was run at 42 °C for 1 h. The reaction was terminated by heating to 70 °C for 10 min.

Fragments of opsin-encoding mRNAs were amplified by the polymerase chain reaction (PCR) using suitable primers. PCR reactions were typically cycled 35 times at 94 °C for 1 min, 48 °C for 1 min and 72 °C for 1 min. Fragments were separated on agarose gels, excised, cloned into t-vectors (Novagen, Madison, WI), and sequenced with the Taq fmol system (Promega, Madison, WI). cDNA molecules encoding one of the opsins eventually identified, MANOP3, were sequenced on an ABI model 373 sequencer using a dye-terminator sequencing kit (Perkin Elmer Applied Biosystems, Foster City, CA). In order to verify the nucleotide sequences of cDNA fragments, 2–4 clones from independent PCR reactions were sequenced on the forward and reverse strands.

For the initial approach to isolating opsin-encoding 3′ cDNA fragments, RT-1 was paired with a degenerate sense-strand primer (OPS-FD, Table 1) designed to a conserved amino acid sequence, (E\D)QAKKMN, from loop 4–5 of known arthropod opsins (Hariyama *et al.* 1993; Smith *et al*. 1993; Towner and Gärtner, 1994; Chang *et al*. 1995). This strategy yielded fragments of two unique opsin-encoding cDNAs, designated MANOP1 and MANOP2 (Chase *et al.* 1996). To amplify upstream regions, antisense primers were designed to unique sequences in MANOP1 (MSN-1, Table 1) and MANOP2 (UVAS-1, Table 1) approximately 150 base pairs downstream from the OPS-FD site. These opsin-specific primers were coupled with a sense degenerate primer (MSH-1, Table 1) designed to a conserved sequence, GNG(MLV)V(IVM)YW, in helix I of insect opsins.

This approach for capturing upstream regions worked only for MANOP1. MANOP1 was then extended into the 5′ untranslated region (UTR) using a 5′ RACE (rapid amplification of cDNA ends) strategy (Frohman *et al*. 1988): cDNA fragments were tailed with dCTP using terminal transferase (Gibco BRL) and reverse transcription reactions were run using a 5′ anchor primer (GIBCO, Table 1) provided with the Gibco kit and a specific antisense primer (LWN-1, Table 1) downstream from the MSH-1 site. The 3′ region of

Table 1. *Forward and reverse primers used for isolating opsin cDNAs*

Forward		
	OPS-FD	CCRTANACRATNGGRTTRTA
	MSH-1	GGNAAYGGNRTNGTNRTNTA
	M3F-1	CCTGTTCATCTGTGCGTGGAC
	GIBCO	GIBCO 5' ANCHOR
Reverse		
	RT-1	CTCTGGGCCCAAGCTTTTTTTTTTTTTTV
	MSN-1	CAGAAGCGGCAAACACCAGCG
	LWN-1	CCACAAGCAAATTTGACGGCGTTT
	UVAS-1	CCATAACGCCGTAGGGTGTCC
	M3R-1	GTCCACGCACAGATGAACAGG
	REV-UV	GCRTAIACRAAIGGRTCDATRCA

```
MANOP1
          M  D   P  G  P   G  L  A   A  L  Q  A   W  A  A   K  S  P   A  Y  G  A   A  N  Q   T  V  V   D  K  V  P
    1 CGGCAATGGA TCCCGGGCCC GGTTTGGCGG CCCTGCAAGC ATGGGCGGCC AAGTCGCCAG CTTACGGCGC TGCAAATCAA ACTGTCGTCG ACAAAGTGCC
      P  D  M   M  H  M   I  D  P  H   W  Y  Q   F  P  P   M  N  P  L   W  H  A   L  L  G   F  T  I  G   V  L  G
  101 TCCAGACATG ATGCACATGA TCGACCCTCA CTGGTATCAA TTTCCGCCTA TGAACCCACT TTGGCATGCG CTGTTAGGAT TTACAATTGG TGTTCTGGGA
      F  V  S   I  S  G  N   G  M  V   I  Y  I   F  M  S  T   K  S  L   K  T  P   S  N  L   V  V  N   L  A  F
  201 TTTGTTTCGA TATCGGGCAA CGGCATGGTC ATCTACATCT TTATGTCAAC TAAGAGTCTT AAAACGCCGT CAAATTTGCT TGTGGTCAAT CTTGCTTTTT
      S  D  F  L   M  M  C   A  M  S   P  A  M  V   V  N  C   Y  Y  E   T  W  V  W   G  P  F   A  C  E   L  Y  A  C
  301 CGGACTTCCT CATGATGTGC GCCATGTCTC CAGCTATGGT AGTTAATTGT TACTATGAAA CTTGGGTATG GGGTCCTTTT GCATGCGAAT TGTACGCCTG
      A  G  S   L  F  G   C  A  S  I   W  T  M   T  M  I   A  F  D  R   Y  N  V   I  V  K   G  I  A  A   K  P  M
  401 TGCTGGTTCT CTATTTGGAT GCGCTTCAAT TTGGACAATG ACTATGATAG CTTTCGACCG ATACAATGTA ATCGTTAAAG GTATAGCAGC TAAGCCCATG
      T  S  N   G  A  L  L   R  I  L   G  I  W   V  F  S  L   A  W  T   L  L  P   F  F  G  W   N  R  Y   V  P  E
  501 ACCAGCAACG GCGCCCTTCT TCGCATACTT GGAATCTGGG TATTTTCACT AGCTTGGACT CTGCTCCCCT TCTTTGGCTG GAACAGATAT GTGCCCGAAG
      G  N  M  T   A  C  G   T  D  Y   L  S  K  S   W  V  S   R  S  Y   I  L  I  Y   S  V  F   V  Y  F   L  P  L  L
  601 GAAACATGAC TGCTTGCGGA ACAGATTACT TGTCCAAGAG TTGGGTCAGC CGAAGCTACA TCCTTATCTA CTCCGTTTTC GTTTACTTCT TGCCGCTTCT
      L  I  I   Y  S  Y   F  F  I  V   Q  A  V   A  A  H   E  K  A  M   R  E  Q   A  K  K   M  N  V  A   S  L  R
  701 TCTTATCATC TATTCTTACT TCTTTATTGT TCAGGCCGTA GCTGCTCACG AAAAGGCAAT GAGGGAACAG GCTAAGAAAA TGAATGTGGC CTCCCTCAGA
      S  S  E   A  A  N  T   S  A  E   C  K  L   A  K  V  A   L  M  T   I  S  L   W  F  M  A   W  T  P   Y  L  V
  801 TCTTCAGAAG CGGCAAACAC CAGCGCTGAA TGCAAACTAG CCAAGGTTGC ATTAATGACC ATTTCCTTAT GGTTCATGGC GTGGACACCA TACCTTGTCA
      I  N  Y  T   G  V  F   E  S  A   P  I  S  P   L  A  T   I  W  G   S  L  F  A   K  A  N   A  V  Y   N  P  I  V
  901 TCAACTATAC CGGAGTGTTT GAAAGCGCAC CCATTAGCCC TCTGGCCACT ATTTGGGGCT CCCTCTTCGC CAAGGCTAAC GCCGTCTACA ATCCTATTGT
      Y  G  I   S  H  P   K  Y  Q  A   A  L  Y   A  K  F   P  S  L  Q   C  Q  S   A  P  E   D  A  G  S   V  A  S
 1001 ATATGGCATC AGCCATCCTA AATACCAAGC CGCCCTGTAC GCCAAGTTTC CATCGCTACA GTGCCAGTCG GCTCCAGAGG ACGCCGGCTC GGTTGCTTCT
      G  T  T   A  V  S  E   E  K  P   A  A
 1101 GGCACCACCG CGGTCTCCGA GGAGAAACCT GCAGCGTGAA TGGACTTTAC AAACAACGAA CGACACTGCA GCCAACTGGA AATACGATTC TGCCGAAGCG
 1201 ATGTTCTTTA TTATTTTTTA TATGCAAACA TTTTACATTT GTACAAGCTG TAAACTCGAA CACACGAGCG TGTGCTGGTT GCAAAATATT CAGGTTAGGT
 1301 CATTACATGT AACATTTAGT ATATGACTCC TGCGTATTTT GAGTTCACGT TTTGTGTCTC GAGTTGAACG TTGCACACTT TTCTTATGCC TTACATATGC
 1401 TTGAAATTAC GGTTAGCGAC ATGAAAAAAA TAAAAAAAAA ATACAGTGAA TATATAATGT CGTTAGACCG TTCCTTTACC TATGAGACTG CTCCCTTAGA
 1501 TATAAATAAC GATAGATTTC TTAACGACGA AATAAGCTCT GTATCTTGGA CAAATAAAGC ATATTATTAA GCAAGCTTGC TGGCTGATGG CCAGAAAAAA
 1601 AAAAAAAAAA

MANOP2
                                                                                                  M  N  N  Q   S  E  N  Y
    1 AGCAGCTCAG TGAACGGAAT GGGCCCAGGA GCACTCTACT GAGGACTGCT CACCCGCAAC CAGAAATTAT TCCCAACATG AACAACCAAT CAGAAAATTA
      Y  H  G   A  Q  F   E  A  L  K   S  A  G   A  I  E   M  L  G  D   G  L  T   G  D  D   L  A  A  I   P  E  H
  101 TTATCATGGC GCACAGTTTG AAGCACTTAA GTCCGCCGGG GCCATTGAAA TGCTCGGAGA TGGTTTGACA GGCGACGATC TGGCAGCGAT ACCGGAGCAC
      W  L  S   Y  P  A  P   P  A  S   A  H  T   A  L  A  L   L  Y  I   F  F  T   F  A  A  L   V  G  N   G  M  V
  201 TGGCTATCGT ACCCGGCGCC GCCCGCTAGT GCTCACACCG CACTGGCACT ACTCTACATT TTCTTTACTT TCGCCGCATT AGTTGGAAAT GGAATGGTCA
      I  F  I  F   S  T  T   K  S  L   R  T  S  S   N  F  L   V  L  N   L  A  I  L   D  F  I   M  M  A   K  A  P  I
  301 TATTCATATT CTCAACGACA AAAAGTTTAC GAACATCAAG CAATTTCCTT GTCTTAAATC TAGCTATACT AGATTTTATA ATGATGGCGA AAGCTCCTAT
      F  I  Y   N  S  A   M  R  G  F   A  V  G   T  V  G   C  Q  I  F   A  L  M   G  A  Y   S  G  I  G   A  G  M
  401 ATTCATCTAT AATAGCGCGA TGCGAGGCTT TGCGGTAGGC ACAGTCGGTT GCCAAATATT CGCTTTGATG GGCGCTTACA GCGGCATCGG AGCTGGCATG
      T  N  A   C  I  A  Y   D  R  H   S  T  I   T  R  P  L   D  G  R   L  S  E   G  K  V  L   L  M  V   A  F  V
  501 ACTAACGCAT GCATAGCTTA TGATAGACAC TCAACTATTA CGCGACCTCT CGATGGAAGA TTGTCGGAAG GCAAGGTCTT ACTCATGGTG GCCTTTGTGT
      W  I  Y  S   T  P  W   A  L  L   P  L  L  K   I  W  G   R  Y  V   P  E  G  Y   L  T  S   C  S  F   D  Y  L  T
  601 GGATATACTC TACGCCTTGG GCTCTCCTTC CACTCTTGAA AATTTGGGGC AGATACGTAC CTGAGGGCTA CCTAACATCG TGTTCTTTCG ACTATCTAAC
      N  T  F   D  T  K   L  F  V  A   C  I  F   T  C  S   Y  V  F  P   M  S  L   I  I  Y   F  Y  S  G   I  V  K
  701 CAACACGTTC GACACGAAGC TGTTCGTAGC GTGTATATTC ACTTGCAGTT ACGTTTTCCC AATGTCCTTA ATAATATATT TCTACAGTGG CATCGTCAAG
      Q  V  F   A  H  E  A   A  L  R   E  Q  A   K  K  M  N   V  E  S   L  R  A   N  Q  G  G   S  S  E   S  A  E
  801 CAGGTGTTTG CTCACGAAGC AGCTTTGAGG GAACAAGCTA AAAAGATGAA CGTAGAATCT CTAAGAGCTA ATCAGGGCGG GTCATCAGAG TCGGCAGAGA
      I  R  I  A   K  A  A   L  T  V   C  F  L  F   V  A  S   W  T  P   Y  G  V  M   A  L  I   G  A  F   G  N  Q  Q
  901 TCAGAATAGC AAAAGCAGCA CTCACTGTTT GCTTTCTGTT CGTGGCCTCG TGGACACCCT ACGGCGTTAT GGCTCTCATA GGCGCGTTTG GAAATCAACA
      L  L  T   P  G  V   T  M  I  P   A  V  A   C  K  A   V  A  C  I   S  P  W   V  Y  A   I  R  H  P   M  Y  R
 1001 ACTTCTTACT CCTGGGGTTA CGATGATTCC TGCAGTGGCG TGTAAAGCTG TAGCGTGTAT TAGTCCTTGG GTATACGCAA TCAGACATCC TATGTACAGG
      Q  E  L   Q  R  R  M   P  W  L   Q  I  D   E  P  D  D   T  V  S   T  A  T   S  N  T  T   N  S  A   P  P  A
 1101 CAAGAGCTGC AACGTCGCAT GCCGTGGTTA CAAATTGATG AGCCCGATGA CACTGTCTCG ACGGCGACCA GCAACACTAC CAACAGCGCG CCACCAGCTG
      A  T  A
 1201 CCACCGCCTA AGTTCTCTAT GCTCAAAAGG GATGACATAT ACTCAGTACA TATACATGTA TTTATTCATT GTGATTTTAA CGACAACTAT GGCTATGCTG
 1301 AAGGCTTTTG ATGTTAAATT ACTGTAGAAT TTATTAAAGA AGCATAAAGT GTCACAGAAC ACCTAAAAAA AAAAAAAAA

MANOP3
    1 ACAAGTGAGA AGCGACTTGG AAGTCGGATA ACTCACTTCG TTCATAAAAA CATCATTTTG TTTTTTTAAC ATCACGATTT GTAACAACTT GTAAATTAAT
                                                                                  M  A  T   N  F  T   Q  E  L   Y  E  I  G
  101 GAAGCTGTAG CAAGTACATG AATTTACTCC GTCGGCCATT TTGTTAACAA CACACGCACA AAATGGCGAC CAATTTCACG CAGGAACTTT ACGAGATCGG
      P  M  A   Y  P  L   K  M  I  S   K  D  V   A  E  H   M  L  G  W   N  I  P   E  E  H   Q  D  L  V   H  D  H
  201 ACCGATGGCC TACCCGCTGA AGATGATATC AAAGGATGTC GCGGAGCACA TGCTGGGTTG GAACATTCCA GAAGAACATC AAGACTTGGT CCACGACCAC
      W  R  N   F  P  A  V   S  K  Y   W  H  Y   V  L  A  L   I  Y  T   M  L  M   V  T  S  L   T  G  N   G  I  V
  301 TGGCGAAACT TCCCCGCGGT CAGCAAGTAC TGGCATTATG TCCTAGCGCT GATATACACC ATGCTCATGG TCACGTCGCT CACCGGCAAT GGAATTGTCA
      I  W  I  F   S  T  S   K  S  L   R  S  A  S   N  M  F   V  I  N   L  A  V  F   D  L  M   M  M  L   E  M  P  L
  401 TTTGGATATT TAGCACTTCA AAATCCCTAC GCAGCGCGAG CAACATGTTC GTGATAAACT TAGCGGTGTT CGATCTGATG ATGATGCTGG AGATGCCGCT
      L  I  M   N  S  F   Y  Q  R  L   V  G  Y   Q  L  G   C  D  V  Y   A  V  L   G  S  L   S  G  I  G   G  A  I
  501 GCTAATCATG AACTCGTTCT ACCAGCGTCT CGTGGGGTAC CAGCTTGGCT GCGATGTGTA CGCGGTACTC GGCTCTCTCT CTGGAATTGG AGGGGCCATC
      T  N  A   V  I  A  F   D  R  Y   K  T  I   S  S  P  L   D  G  R   I  N  T   V  Q  A  G   L  L  I   A  F  T
  601 ACTAATGCTG TCATTGCTTT TGATCGATAC AAAACAATTT CGTCGCCACT TGATGGAAGA ATAAACACAG TACAAGCGGG TCTGCTGATC GCCTTCACCT
      W  F  W  A   L  P  F   T  I  L   P  A  F  R   I  W  G   R  F  V   P  E  G  F   L  T  T   C  S  F   D  Y  F  T
  701 GGTTCTGGGC GTTGCCGTTC ACTATCCTGC CGGCGTTCAG GATATGGGGG AGATTTGTTC CCGAGGGCTT CCTGACGACG TGCTCGTTCG ACTATTTCAC
      E  D  Q   D  T  E   V  F  V  A   C  I  F   V  W  S   Y  C  I  P   M  A  L   I  C  Y   F  Y  S  Q   L  F  G
  801 GGAGGACCAG GATACGGAGG TATTTGTCGC GTGCATCTTC GTGTGGAGCT ACTGCATACC GATGGCGCTC ATCTGTTACT TCTACTCTCA ACTATTCGGT
      A  V  R   L  H  E  R   M  L  Q   E  Q  A   K  K  M  N   V  K  S   L  A  S   N  K  E  D   N  S  R   S  V  E
  901 GCTGTCCGCC TCCATGAACG CATGCTACAA GAACAAGCCA AGAAGATGAA CGTGAAATCA TTAGCATCAA ACAAGGAAGA CAACAGCCGC AGCGTAGAGA
      I  R  I  A   K  V  A   F  T  I   F  F  L  F   I  C  A   W  T  P   Y  A  F  V   T  M  T   G  A  F   G  D  R  T
 1001 TCAGGATAGC TAAAGTGGCG TTCACTATTT TCTTCCTGTT CATCTGTGCG TGGACTCCTT ACGCTTTCGT CACCATGACA GGCGCATTTG GCGACAGGAC
      L  L  T   P  I  A   T  M  I  P   A  V  C   C  K  V   V  S  C  I   D  P  W   V  Y  A   I  N  H  P   R  Y  R
 1101 CCTACTAACT CCGATAGCAA CGATGATCCC AGCCGTGTGC TGTAAAGTAG TTTCGTGCAT AGACCCATGG GTGTACGCCA TCAACCATCC GAGATACAGG
      A  E  L   Q  K  R  L   P  W  M   G  V  R   E  Q  D  P   D  A  V   S  T  T   T  S  V  A   T  A  G   F  Q  P
 1201 GCGGAGTTGC AGAAGCGTCT GCCATGGATG GGCGTCCGCG AGCAAGACCC TGACGCGGTA TCTACCACCA CCAGCGTCGC CACCGCTGGT TTTCAACCAC
      P  A  A  E   A
 1301 CAGCCGCGGA GGCCTAGACA GACAAACTGA TTCATAATTA CTGAGGAAAA TCATGTTAAA GCTTAAGCAA TAATTTAAAA CTCCATTAAG CAAAAAAAAA
 1401 AAAAAAA
```

Appendix. Sequences of opsin-encoding cDNAs from the retina of *Manduca sexta* with amino acid translations. The sequences have been deposited in the Genome Sequence Database (www.ncgr.org/gsdb) under the following accession numbers: MANOP1, 76082; MANOP2, 109852; MANOP3, 1249561.

MANOP2 was isolated directly using the 5′ RACE approach by pairing the GIBCO primer with the UVAS-1 primer.

Fragments of a third opsin-encoding cDNA, MANOP3, were generated using OPS-FD and another reverse primer (REV-UV, Table 1) designed to a nucleotide sequence in helix VI of a honeybee opsin (Chang, 1995; see also Bellingham *et al*. 1997). Downstream fragments were obtained by pairing RT-1 with a primer (M3F-1, Table 1) specific to a region downstream from the OPS-FD site. Upstream regions were obtained using the 5′ RACE approach, pairing another MANOP3-specific primer (M3R-1) with the GIBCO primer.

Phylogenetic analysis

Conceptual translations of the opsin-encoding cDNAs from *Manduca sexta* were compared, using FASTA (Pearson, 1990), with other known arthropod opsins deposited in GenBank and aligned using CLUSTAL W (Thompson *et al*. 1994). Phylogenetic trees were constructed with PAUP (Swofford, 1993) using maximum parsimony and neighbor-joining methods. All trees were bootstrapped a minimum of 100 times to estimate confidence intervals.

Results

Opsin-encoding cDNAs from the retina of *Manduca sexta* were initially targeted with a degenerate primer (OPS-FD, Table 1) designed to an amino acid sequence, (E\D)QAKKMN, that is conserved in arthropod opsins in the loop between transmembrane helices V and VI (Hariyama *et al.* 1993; Smith *et al.* 1993; Towner and Gärtner, 1994). Among the many 3′ fragments that were isolated by this approach, we identified two distinct opsin cDNAs, designated MANOP1 and MANOP2 (Chase *et al.* 1996).

Although three rhodopsins had been extracted from *Manduca sexta* retinas (Bennett and Brown, 1985), efforts to isolate a third opsin-encoding cDNA with the 3′ RACE technique were unsuccessful. While our work was in progress, Chang (1995) isolated fragments of three opsin-encoding cDNAs from the honeybee (*Apis mellifera*). Two of the honeybee cDNAs were very similar, respectively, to MANOP1 and MANOP2; the third (whose full sequence was recently published by Bellingham *et al.* 1997) was somewhat different. We isolated the third *Manduca sexta* cDNA, MANOP3, using a primer (REV-UV, Table 1) based on a sequence in helix VI of the third honeybee cDNA. The three opsin cDNAs isolated from *Manduca sexta* retinas and their translations are shown in the Appendix. The amino acid sequences of the opsins are compared in Fig. 1.

MANOP1 cDNA contains 1594 base pairs (bp) with a single open reading frame encoding an opsin with 377 amino acids. MANOP2 (1362 bp) also encodes 377 residues, while

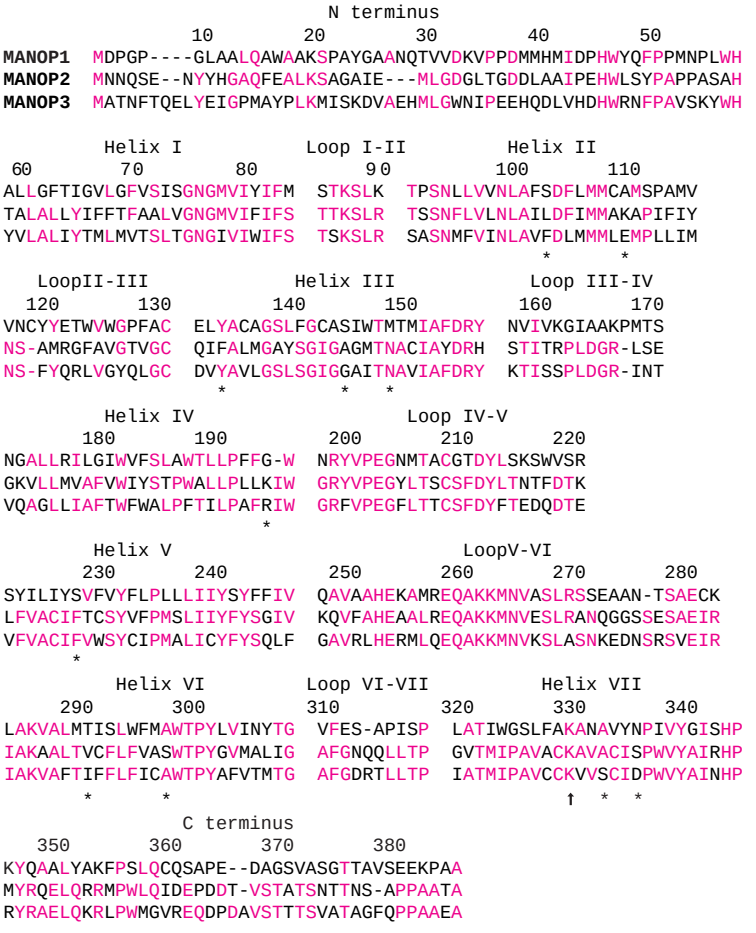

Fig. 1. Aligned amino acid sequences of three opsins isolated as cDNA from the retina of *Manduca sexta*. M1, MANOP1; M2, MANOP2; M3, MANOP3. See Appendix for nucleotide sequences. Transmembrane α-helical segments are defined according to Baldwin's (1993) model of G-protein-coupled receptors. Residues identical in two or all three of the sequences are in red. The arrow indicates the site of the chromophore Schiff-base linkage. Asterisks indicate amino acid positions listed in Table 3.

MANOP3 (1391 bp) encodes 384 residues. The approximate sizes of the corresponding opsin-encoding mRNAs were measured on northern blots of retinal extracts: MANOP1 mRNA ≈1.8 kb, MANOP2 mRNA ≈1.5 kb and MANOP3 mRNA ≈2.2 kb (Chase *et al.* 1996; Landry, 1997).

Several considerations indicate that the we have not isolated the full 3′ untranslated sequences of MANOP2 and MANOP3. Their 3′ UTRs are short in comparison with that of MANOP1 and they lack a polyadenylation signal: AATAAA (Rice *et al.* 1992). Furthermore, MANOP3 mRNA is almost twice as large as the cDNA we have isolated. It is likely that the RT-1 primer (Table 1) found stretches of adenines in the 3′ UTRs of MANOP2 and MANOP3 rather than annealing to the terminal poly(A) tail. Such internal adenine repeats are present between nucleotides 1424 and 1441 in the 3′ UTR of MANOP1. Secondary structure arising from a GC-rich region prevented sequencing far into the 5′ UTR of MANOP1.

The deduced amino acid sequences of the three *Manduca sexta* opsins show the various conserved characteristics of opsins generally and of arthropod opsins in particular (Chang *et al.* 1995; Gärtner and Towner, 1995; Yokoyama, 1995; Popp *et al.* 1996). Sequence similarities and hydropathy profiles (not shown) confirm the seven membrane-spanning helical domains that characterize opsins and other G-protein-coupled receptors (Baldwin, 1993). A lysine (Lys-330) residue in the middle of helix VI is the site of attachment for the Schiff-base linkage of the retinaldehyde chromophore. Cysteine residues at the extracellular end of helix III (Cys-131) and within extracellular loop IV–V (Cys-209) presumably provide a stabilizing disulfide linkage (Gärtner and Towner, 1995). Asparagine residues susceptible to glycosylation are present in the N-terminal region (MANOP1, Asn-28; MANOP2, Asn-3; MANOP3, Asn-4); as in some other invertebrate opsins (Popp *et al.* 1996), there is a second potential glycosylation site (Asn-205) in extracellular loop IV–V of MANOP1. There are a number of serine/threonine potential phosphorylation sites in the C termini. Conserved sequences, particularly those that interact with G-protein (transducin), are present in the helix-connecting loops that extend beyond the membrane surfaces (Gärtner and Towner, 1995; Yokoyama, 1995).

Of the amino acids, 25 % are identical in all three opsins of *Manduca sexta*. One-third of the residues are found to be the same when MANOP1 is compared with either MANOP2 or MANOP3. MANOP2 and MANOP3 are more closely related, showing 46 % identity (Fig. 1; Table 2).

Discussion

Opsins and rhodopsins

We isolated opsin-encoding cDNAs from the retina of *Manduca sexta* in order to characterize further the three rhodopsins of the compound eye: P520, P450 and P357. The next step is to decide which opsin forms which rhodopsin. It is likely that MANOP1 is the opsin of P520: MANOP1 mRNA is 10 times more abundant than the other two opsin mRNAs (Chase *et al.* 1996; Landry, 1997), and P520, to a similar degree, is the most plentiful photopigment in the retina (Bennett and Brown, 1985; Bennett *et al.* 1997). We have no equivalent evidence for deciding between P450 and P357 in the assignment of MANOP2 and MANOP3.

Comparison of the hawkmoth opsins with those of other arthropods offers another avenue to their identification (Fig. 2; Tables 2, 3). The phylogenetic topography of arthropod opsins is growing clearer as more sequences are collected from diverse taxa and the different spectral classes of photoreceptors. Three major divisions are evident in the phylogenetic tree of arthropod opsins (Fig. 2): branch A, opsins from the crab *Hemigrapsis sanguineus*; branch B, a

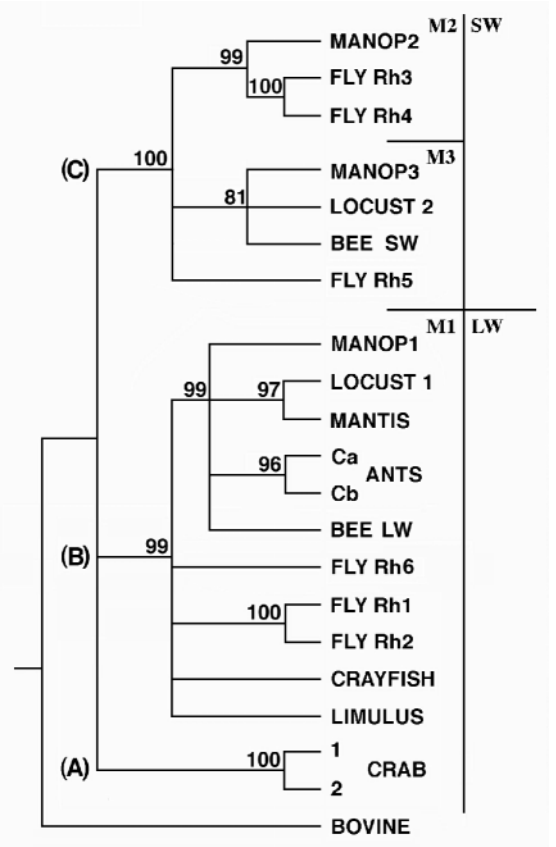

Fig. 2. Phylogenetic tree for arthropod opsins using bovine opsin as outgroup. Numbers above nodes represent bootstrap percentages of 100 replications. (A,B,C) Major phyletic branches; LW, putative long-wavelength-sensitive rhodopsins; SW, putative short-wavelength-sensitive rhodopsins; M1, M2, M3, proposed functional groups designated according to the inclusion of MANOP1, MANOP2 or MANOP3. MANOP1, 2, 3, *Manduca sexta.* MANTIS: *Sphodromantis* sp. (Towner and Gärtner, 1994). ANTS: Cb, *Cataglyphis bombycina*; Ca, *Camponotus abdominalis* (Popp *et al.* 1996). LOCUST 1,2: *Schistocerca gregaria* (Towner *et al.* 1997). FLY: *Drosophila melanogaster* (Pollock and Benzer, 1988; Chou *et al.* 1996; Huber *et al.* 1997); similar opsins from other flies are not included. BEE: *Apis mellifera* (Chang *et al.* 1996; Bellingham *et al.* 1997). LIMULUS: *Limulus polyphemus*, two nearly identical opsins from the lateral compound eye and ocellus (Smith *et al.* 1993). CRAYFISH: *Procambarus clarkii* (Hariyama *et al.* 1993) CRAB 1,2: *Hemigrapsus sanguineus* (Sakamoto *et al.* 1996).

large group that includes opsins from *Limulus polyphemus*, the crayfish *Procambarus clarkii* and a number of insects; and branch C, another cluster of insect opsins.

Several of the opsins in branch B have been be identified with rhodopsins absorbing from approximately 480 nm (Rh1 and Rh2 fly opsins: O'Tousa *et al.* 1985; Cowman *et al.* 1986; Huber *et al.* 1990) to 520 nm (*Limulus polyphemus* opsins: Smith *et al.* 1993). We therefore classify these opsins as belonging to an 'LW' group of functionally similar arthropod photopigments tuned for absorbance at long wavelengths. Similarities at key amino acid positions identified in Table 3 and spectral sensitivity measurements from the crab eye suggest that the opsins of *Hemigrapsus sanguineus* should also be included among the LW opsins. The membership of MANOP1 in the LW group is further support for its identification as P520. It is found in a cluster (designated M1) of opsins from orthopteran and hymenopteran species with amino acid identities of 70 % or greater (Table 2).

Branch C in Fig. 2 includes, at present, only insect opsins: MANOP2 and MANOP3, the Rh3, Rh4 and Rh5 opsins of flies (Montell *et al.* 1987; Pollock and Benzer, 1988; Chou *et al.* 1996) and recently characterized opsins from the honeybee (Bellingham *et al.* 1997) and a locust (Towner *et al.* 1997). As the only identified rhodopsins of this group are the ultraviolet-absorbing photopigments of *Drosophila* species, Rh3 and Rh4, we designate this the short-wavelength (SW) cluster of insect opsins (Fig. 2; Table 2). It will be interesting to determine whether short-wavelength rhodopsins from arthropods other than insects will join this phyletic group. As both MANOP2 and MANOP3 are situated in the SW group, we infer that the SW division includes both ultraviolet- and blue-absorbing rhodopsins. Certain ultraviolet-, violet- and blue-sensitive photopigments of vertebrate cones similarly constitute a short-wavelength 'tritope-group' of related sequences (Hisatomi *et al.* 1996; Yokoyama, 1995; Yokoyama and Yokoyama, 1996). Since MANOP2 is more like the known ultraviolet opsins of flies, we suggest that MANOP2 encodes P357 whereas MANOP3 encodes P450.

Apart from the close similarity of Rh3 and Rh4, the SW opsins are structurally diverse, with amino acid identities ranging from 46 % to 62 % (Table 2). For purposes of comparison (with no necessary functional implication at this point), we have distinguished M2 and M3 subgroups (Fig. 2; Table 2). The M2 group includes only the ultraviolet-sensitive fly rhodopsins and MANOP2. The more disparate M3 group includes, in addition to MANOP3, locust Lo-2, a honeybee opsin and Rh5 of *Drosophila melanogaster*.

Opsins of the M3 group have been assigned by a variety of criteria. There is indirect evidence that Rh5 encodes a blue-sensitive rhodopsin (Chou *et al.* 1996). Gärtner and Towner (1995) have suggested that Lo-2, one of two cDNAs isolated from the retina of *Schistocerca gregaria*, may be the opsin of a blue-sensitive photopigment. The existence of two rhodopsins in the retina of another orthopteran species, *Locusta migratoria*, one absorbing in the green (λ_{max} 500–515 nm) and

Table 2. *Percentage residue identities among amino acid sequences of selected arthropod opsins*

		FLY Rh5	BEE SW	LOCUST 2	MANOP3	FLY Rh3	MANOP2	CRAB Rh 2	CRAYFISH	LIMULUS	FLY Rh6	FLY Rh1	BEE LW	LOCUST 1	ANT Cb	MANTIS	MANOP 1
Long-wavelength M1	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	
MANOP1	1	33.2	31.4	33.8	32.7	37.6	33.3	41.2	56.2	60.6	58.5	62.3	71.2	74.5	75.5	76.9	100
MANTIS	2	34.0	32.6	34.9	34.5	38.9	34.2	41.0	56.1	60.3	57.8	62.6	73.7	85.6	77.7	100	
ANT Cb	3	35.0	32.8	37.3	36.3	39.2	34.7	41.2	54.8	60.6	57.5	63.9	78.1	77.4	100		
LOCUST 1	4	33.9	33.5	34.5	34.8	38.6	34.0	39.7	54.5	59.4	59.8	61.4	70.1	100			
BEE LW	5	32.4	31.5	35.7	34.2	37.0	34.4	38.3	55.9	54.6	56.7	63.2	100.0				
FLY Rh1	6	35.4	36.8	38.4	37.6	38.9	38.2	44.1	55.2	55.5	53.7	100					
FLY Rh6	7	29.8	31.3	32.4	31.6	33.5	30.8	36.8	48.5	50.1	100.0						
LIMULUS	8	31.8	35.0	34.8	33.9	41.1	34.6	40.4	56.2	100							
CRAYFISH	9	32.0	30.9	30.9	31.7	37.4	32.2	41.7	100								
CRAB Rh2	10	36.6	34.0	34.3	34.3	40.3	34.9	100									
Short-wavelength M2																	
MANOP2	11	46.0	45.6	44.9	46.4	62.4	100										
FLY Rh3	12	46.0	46.1	47.5	50.9	100											
Short-wavelength M3																	
MANOP3	13	49.7	56.5	59.0	100												
LOCUST 2	14	46.2	60.8	100													
BEE SW	15	48.9	100														
FLY Rh5	16	100															

See legend to Fig. 2 for species identification and the basis for grouping insect opsins into long-wavelength and short-wavelength.

Table 3. *Amino acid positions in transmembrane α-helices that may be significant for spectral tuning in insect rhodopsins*

	Helix II		Helix III			Helix IV		Helix V	Helix VI		Helix VII	
Manduca no.	103	110	134	145	149	195	198	228	291	298	333	336
Baldwin no.	13	21	3	14	18	24	27	7	8	15	14	17
Long-wavelength M1												
MANOP1	Ser (p)	Ala	Tyr (p)	Ser (p)	Met	Gly		Ser (p)	Thr (p)	Ala	Ala	Asn (p)
MANTIS	*	Ser (p)	*	*	*	*		Ser(p)	Thr (p)	Ala	*	*
LOCUST 1	*	Thr (p)	*	*	*	*		Ser(p)	Thr (p)	Gly	*	*
BEE LW	*	Cys	*	*	*	*		Gly	Thr (p)	Ala	*	*
ANT Ca	*	Cys*	*	*	*	*		Ser(p)	Thr (p)	Ala	*	*
ANT Cb	*	Ser (p)	*	*	*	*		Ser(p)	Thr (p)	Ala	*	*
FLY Rh1	*	Thr (p)	*	*	*	*		Ser(p)	Thr (p)	Ala	*	*
FLY Rh2	*	Ser(p)	*	*	*	*		Ser(p)	Thr (p)	Ala	*	*
FLY Rh6	*	Thr(p)	*	*	*	*		Ser(p)	Glu(-)	Ala	*	*
Short-wavelength M2												
MANOP2	Leu	Lys (+)	Phe	Ala	Asn (p)	Lys (+)		Phe	Val	Ser (p)	Ala	Ser (p)
FLY Rh3	Cys	*	*	*	*	Glu (–)		*	Ile	*	*	Asp (–)
FLY Rh4	Phe	*	*	*	*	Gln (p)	Asp (–)	*	Ile	*	*	Asp (–)
Short-wavelength M3												
MANOP3	Phe	Glu (–)	Tyr (p)	Gly	Asn (p)	Arg (+)		Phe	Ile	Ala	Ser (p)	Asp (–)
FLY Rh5	*	Asn (p)	*	Ala	*	Gln (p)		*	*	Ala	*	*
LOCUST 2	*	Glu (–)	*	Ser (p)	*	Arg (+)		*	*	Ser (p)	*	*
BEE SW	*	Glu (–)	*	Gln (p)	*	Lys (+)		*	*	Ala	*	*
	▲					▲	▲					▲

Positions are numbered as in Fig. 1 (*Manduca* numbers), and in each helix according to the model of Baldwin (1933). Most face inward toward the chromophore according to the Baldwin model. Arrowheads indicate positions significant to the external point-charge model of spectral tuning proposed by Towner *et al.* (1997); position 198 is pertinent only to Rh4. Asterisks indicate that all residues are identical within a group; (p) polar residues, (+) basic residues, (–) acidic residues.

See Fig. 2 for species identification and the basis for grouping insect opsins into long-wavelength and short-wavelength M2 and M3 categories.

another in the blue (λ_{max} 450–430 nm) region, has been inferred from intracellular spectral sensitivity measurements (Bennett *et al.* 1967; Lillywhite, 1978). Gärtner and Towner (1995) suggested that Lo-1 and Lo-2 of *Schistocerca gregaria* are, respectively, the opsins of similar photopigments. However, it would be surprising if locusts lacked ultraviolet receptors, and there is indirect evidence for their existence (Osorio, 1986). Thus, the identification of Lo-2 with a blue- rather than an ultraviolet-sensitive rhodopsin is plausible but not compelling.

Chang (1995) isolated partial sequences of three honeybee opsins that are similar to the three opsins of *Manduca sexta*. She proposed that the bee homologue of MANOP2 forms its ultraviolet-sensitive pigment and identified the homologue of MANOP3 with the blue-sensitive bee rhodopsin. The latter has since been fully characterized by Bellingham *et al.* (1997); although they describe it, contrary to Chang's hypothesis, as the ultraviolet-sensitive rhodopsin, it has not been empirically identified.

Spectral tuning of photopigments

The regulation of rhodopsin absorbance by opsin structure – presumably through the distribution of charged and polar amino acid residues in the binding pocket around the chromophore – is a challenging problem of current interest. The basis of blue-shifted spectra is particularly uncertain. Recent models and crystallographic data (Baldwin, 1993; Schertler *et al.* 1993; Schertler and Hargrave 1995; Davies *et al.* 1996) provide a foundation for theoretical, comparative and phylogenetic approaches to the problem (Chang *et al.* 1995; Gärtner and Towner, 1995; Yokoyama, 1995). Attention has focused on transmembrane helices III, VI and VII (Schertler *et al.* 1993; Schertler and Hargrave 1995; Davies *et al.* 1996). Amino acid residues of these helices have been shown experimentally to interact with the chromophore in vertebrate rhodopsins (Yokoyama, 1995). However, the details of spectral tuning in vertebrate rhodopsins (Lin *et al.* 1992; Hunt *et al.* 1996; Yokoyama, 1995) seem to hold little pertinence for understanding the properties of invertebrate photopigments (Britt *et al.* 1993; Chang *et al.* 1995; Gärtner and Towner, 1995). It appears that the assortment of photopigments underlying color vision has evolved independently in insects and vertebrates through distinct opsin/chromophore interactions (Fryxell and Myerowitz, 1991; Chang *et al.* 1995; Gärtner and Towner, 1995).

A number of amino acid positions have been suggested as significant for spectral tuning in invertebrate rhodopsins. For the following discussion, we use the '*Manduca* numbers' of

Fig. 1 as well as the system of Baldwin (1993), who numbered the residues in each helix, according to her model, from 1 to 26 (Table 3). The inward-facing side of helix III is positioned close to the Schiff-base linkage of the chromophore at Lys-330 in helix VI. In vertebrate but not invertebrate photopigments, a highly conserved glutamate residue near the intradiscal (extracellular) end of helix III functions as the negative counterion that stabilizes the protonated Schiff base (Sakmar *et al.* 1989, 1991; Lin *et al.* 1992). There is a polar tyrosine residue at that position (Tyr-134) in most invertebrate rhodopsins, a residue that may play a role in long-wavelength tuning (Hall *et al.* 1991; Smith *et al.* 1993; Chang *et al.* 1995; Gärtner and Towner, 1995; Popp *et al.* 1996) since it is replaced by non-polar phenylalanine in the ultraviolet-sensitive rhodopsins of flies (Table 3). Position 134 is also occupied by phenylalanine in MANOP2, whereas MANOP3 and the other members of the M3 group resemble the LW opsins in having tyrosine at that position (Table 3). Chang *et al.* (1995) focused on another residue in helix III, suggesting that polar Ser-145 cooperates with Tyr-134 in stabilizing the protonated Schiff base in long-wavelength rhodopsins; non-polar residues occupy that position in most members of the M2 and M3 groups.

Bellingham *et al.* (1997) catalogued five sites in arthropod opsins, all inward-facing according to the model of Baldwin (1993), which consistently differed, with regard to polarity or charge, between long- and short-wavelength-sensitive rhodopsins: positions 103 in helix II, 149 in helix III, 228 in helix V, 291 in helix VI and 336 in helix VII. Table 3 shows that these generalizations hold up well in differentiating LW from SW insect rhodopsins.

Towner *et al.* (1997) have proposed an external two-point charge model for short-wavelength-shifted insect rhodopsins on the basis of a negatively charged counterion in helix VII (position 336) and a positive/negative pair of charged residues located in helices II (position 110) and IV (position 195 or, in Rh4, position 198) that form a dipole. However, MANOP2 is inconsistent with this model. Position 336 is occupied by a polar residue in MANOP2, as it is in the long-wavelength-absorbing rhodopsins of the M1 group. Furthermore, positions 110 and 195 are both occupied by lysine, rather than a positive/negative pair of residues. *Drosophila melanogaster* Rh5 lacks charged residues at the proposed dipole positions.

Although the separation of insect SW opsins into M2 and M3 groups is not strongly supported by overall sequence comparison, there are several consistent differences among residues facing inward around the chromophore. Most notable is polar tyrosine residue at position 134 in the M3 opsins, since it may play a crucial role in shifting absorbance to longer wavelengths (Hall *et al.* 1991). Other sites that consistently differ with regard to residue attributes between the M2 and M3 opsins are position 110 in helix II and position 333 in helix VII. We propose, in accordance with the speculations of others (Gärtner and Towner, 1995; Chang, 1995; Chou *et al.* 1996; Towner *et al.* 1997), that the rhodopsins of the M3 group represent a new class of blue-sensitive insect rhodopsins.

We clearly need direct evidence associating opsin sequences with rhodopsin spectra in insect species other than flies. It is a distinct advantage to have the sequences for a complete set of opsins that are expressed in a retina containing the three typical rhodopsins of insect compound eyes. We know enough about the localization of the three characterized *Manduca sexta* photopigments, both at the cellular level (green- and ultraviolet-sensitive cells occupy different positions in the retinula: Cutler *et al.* 1995) and across the retina (blue-sensitivity has been localized to the ventral half of the retina: Bennett *et al.* 1997) to identify our three opsins by *in situ* hybridization and immunocytochemistry (Xu, 1996).

This work was supported by National Science Foundation Grant IBN-9210933. We thank Kenneth Kleene and Rachel Skvirsky for helpful advice, Richard Kesseli for the use of laboratory facilities, and Belinda Chang for useful discussions and a primer. Lily Landry, Huihong Xu and Amy Tucker provided technical assistance.

References

BALDWIN, J. M. (1993). The probable arrangement of the helices in G protein-coupled receptors. *EMBO J.* **12**, 1693–1703.

BELLINGHAM, J., WILKIE, S. E., MORRIS, A. G., BOWMAKER, J. K. AND HUNT, D. M. (1997). Characterization of the ultraviolet-sensitive opsin gene of the honey bee, *Apis mellifera*. *Eur. J. Biochem.* **243**, 775–781.

BENNETT, R. R. AND BROWN, P. K. (1985). Properties of the visual pigments of the moth *Manduca sexta* and the effects of two detergents, digitonin and CHAPS. *Vision Res.* **25**, 1771–1781.

BENNETT, R. R., TUNSTALL, J. AND HORRIDGE, G. A. (1967). Spectral sensitivity of single retinula cells of the locust. *Z. vergl. Physiol.* **55**, 195–206.

BENNETT, R. R. AND WHITE, R. H. (1989). Influence of carotenoid deficiency on visual sensitivity, visual pigment and P-face particles of photoreceptor membrane in the moth *Manduca sexta*. *J. comp. Physiol.* A **164**, 321–331.

BENNETT, R. R., WHITE, R. H. AND MEADOWS, J. (1997). Regional specialization in the eye of the sphingid moth *Manduca sexta*: blue-sensitivity of the ventral retina. *Vis. Neurosci.* **14**, 523–526.

BERNARD, G. D. AND REMINGTON, C. L. (1991). Color vision in *Lycaena* butterflies: spectral tuning of receptor arrays in relation to behavioral ecology. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2783–2787.

BRITT, S. G., FEILER, R., KIRSCHFELD, K. AND ZUKER, C. S. (1993). Spectral tuning of rhodopsin and metarhodopsin *in vivo*. *Neuron* **11**, 29–39.

CARULLI, J. P., CHEN, D.-M., STARK, W. S. AND HARTL, D. L. (1994). Phylogeny and physiology of *Drosophila* opsins. *J. molec. Evol.* **38**, 250–262.

CHANG, B. S. W. (1995). Opsin phylogeny and evolution. PhD dissertation, Harvard University, Cambridge, MA, USA.

CHANG, B. S. W., AYERS, D., SMITH, W. C. AND PIERCE, N. E. (1996). Cloning of the gene encoding honeybee long-wavelength rhodopsin: a new class of insect long-wavelength visual pigments. *Gene* **173**, 215–219.

CHANG, B. S. W., CRANDALL, K. A., CARULLI, J. P. AND HARTL, D. L. (1995). Opsin phylogeny and evolution: a model for blue shifts in wavelength regulation. *Molec. phylogenet. Evol.* **4**, 31–43.

CHASE, M. R., BENNETT, R. R. AND WHITE, R. H. (1996). Expression of opsin mRNA in normal and vitamin A deficient retinas of the sphingid moth *Manduca sexta*. *Vis. Neurosci.* **13**, 353–358.

CHOU, W.-H., HALL, K. J., WILSON, D. B., WIDEMAN, C. L., TOWNSON, S. M., CHADWELL, L. V. AND BRITT, S. G. (1996). Identification of a novel *Drosophila* opsin reveals specific patterning of the R7 and R8 photoreceptor cells. *Neuron* **17**, 1101–1115.

COWMAN, A. F., ZUKER, C. S. AND RUBIN, G. M. (1986). An opsin gene expressed in only one photoreceptor cell type of the *Drosophila* eye. *Cell* **44**, 705–710.

CUTLER, D. E., BENNETT, R. R., STEVENSON, R. D. AND WHITE, R. H. (1995). Feeding behavior in the nocturnal moth *Manduca sexta* is mediated mainly by blue receptors, but where are they located in the retina? *J. exp. Biol.* **198**, 1909–1917.

DAVIES, A., SCHERTLER, G. F. X., GOWEN, B. E. AND SAIBIL, H. R. (1996). Projection structure of an invertebrate rhodopsin. *J. struct. Biol.* **117**, 36–44.

FROHMAN, M. A., DUSH, M. K. AND MARTIN, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8998–9002.

FRYXELL, K. J. AND MYEROWITZ, E. M. (1991). The evolution of rhodopsins and neurotransmitter receptors. *J. molec. Evol.* **33**, 367–378.

GÄRTNER, W. AND TOWNER, P. (1995). Invertebrate visual pigments. *Photochem. Photobiol.* **62**, 1–16.

GOLDSMITH, T. H. (1990). Optimization, constraint and history in the evolution of eyes. *Q. Rev. Biol.* **65**, 281–322.

HALL, M. D., HOON, M. A., RYBA, N. J. P., POTTINGER, J. D. D., KEEN, J. N., SAIBIL, H. R. AND FINDLAY, J. B. C. (1991). Molecular cloning and primary structure of squid (*Loligo forbesi*) rhodopsin, a phospholipase C-directed G-protein-linked receptor. *Biochem. J.* **274**, 35–40.

HARIYAMA, T., OZAKI, K., TOKUNAGA, F. AND TSUKAHARA, Y. (1993). Primary structure of crayfish visual pigment deduced from cDNA. *FEBS Lett.* **315**, 287–292.

HISATOMI, O., SATOH, T., BARTHEL, L. K., STENKAMP, D. L., RAYMOND, P. A. AND TOKUNAGA, F. (1996). Molecular cloning and characterization of the putative ultraviolet-sensitive visual pigment of goldfish. *Vision Res.* **36**, 933–939.

HUBER, A., SCHULZ, S., BENTROP, J., GROELL, C., WOLFRUM, U. AND PAULSEN, R. (1997). Molecular cloning of *Drosophila* Rh6 rhodopsin: the visual pigment of a subset of R8 photoreceptor cells. *FEBS Lett.* **406**, 6–10.

HUBER, A., SMITH, D. P., ZUKER, C. S. AND PAULSEN, R. (1990). Opsin of *Calliphora* peripheral photoreceptors R1–6. *J. biol. Chem.* **265**, 17906–17910.

HUNT, D. M., FITZGIBBON, J., SLOBODYANYUK, S. J. AND BOWMAKER, J. K. (1996). Spectral tuning and molecular evolution of rod visual pigments in the species flock of cottid fish in Lake Baikal. *Vision Res.* **36**, 1217–1224.

LANDRY, L. H. (1997). Diurnal and circadian control of opsin mrna in the retina of *Manduca sexta*. MS thesis, University of Massachusetts Boston, Boston, MA, USA.

LILLYWHITE, P. G. (1978). Coupling between locust photoreceptors revealed by a study of quantum bumps. *J. comp. Physiol.* A **125**, 13–27.

LIN, S. W., SAKMAR, T. P., FRANKE, R. R., KHORANA, H. G. AND MATHIES, R. A. (1992). Resonance Raman microprobe spectroscopy of rhodopsin mutants: effect of substitutions in the third transmembrane helix. *Biochemistry* **31**, 5105–5111.

MACDONALD, R. J., SWIFT, G. H., PRZYBYLA, A. E. AND CHIRGWIN, J. M. (1987). Isolation of RNA using guanidinium salts. In *Methods in Enzymology*, vol. 152 (ed. S. L. Berger and A. R. Kimmel), pp. 219–227. San Diego: Academic Press.

MENZEL, R. AND BACKHAUS, W. (1991). Colour vision in insects. In*Vision and Visual Dysfunction*, chapter 14, vol. 6 (ed. P. Gouras), pp. 262–293. Houndsmills: Macmillan Press.

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

OSORIO, D. (1986). Ultraviolet sensitivity and spectral opponency in the locust. *J. exp. Biol.* **122**, 193–208.

O'TOUSA, J. E., BAEHR, W., MARTIN, R. L., HIRSH, J., PAK, W. L. AND APPLEBURY, M. L. (1985). The Drosophila *ninaE* gene encodes an opsin. *Cell* **40**, 839–850.

PEARSON, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. In *Methods in Enzymology*, vol. 183 (ed. R. F. Doolittle), pp. 63–98. San Diego: Academic Press.

PEITSCH, D., FIETZ, A., HERTEL, H., DE SOUZA, J., VENTURA, D. F. AND MENZEL, R. (1992). The spectral input systems of hymenopteran insects and their receptor-based colour vision. *J. comp. Physiol.* A **170**, 23–40.

POLLOCK, J. A. AND BENZER, S. (1988). Transcript localization of four opsin genes in the three visual organs of *Drosophila*; RH2 is ocellus specific. *Nature* **333**, 779–782.

POPP, M. P., GRISSHAMMER, R., HARGRAVE, P. A. AND SMITH, W. C. (1996). Ant opsins: sequences from the Saharan silver ant and the carpenter ant. *Invert. Neurosci.* **1**, 323–329.

RICE, P. M., ELLISTON, K. AND GRIBSKOV, M. (1992). DNA. In *Sequence Analysis Primer*, chapter 1 (ed. M. Gribskov and J. Devereux), pp. 1–59. New York: W. H. Freeman.

SAKAMOTO, K., HISATOMI, O., TOKUNAGA, F. AND EGUCHI, E. (1996). Two opsins from the compound eye of the crab *Hemigrapsus sanguineus*. *J. exp. Biol.* **199**, 441–450.

SAKMAR, T. P., FRANKE, R. R. AND KHORANA, H. G. (1989). Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8309–8313.

SAKMAR, T. P., FRANKE, R. R. AND KHORANA, H. G. (1991). The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pK_a. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3079–3083.

SCHERTLER, G. F. X. AND HARGRAVE, P. A. (1995). Projection structure of frog rhodopsin in two crystal forms. *Proc. natn. Acad. Sci. U.S.A.* **92**, 11578–11582.

SCHERTLER, G. F. X., VILLA, C. AND HENDERSON, R. (1993). Projection structure of rhodopsin. *Nature* **362**, 770–772.

SMITH, W. C., PRICE, D. A., GREENBERG, R. M. AND BATTELLE, B.-A. (1993). Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus polyphemus*. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6150–6154.

SWOFFORD, D. L. (1993). *PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1*. Illinois Natural History Survey, Champaign, IL, USA.

THOMPSON, J. D., HIGGINS, D. G. AND 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.

TOWNER, P. AND GÄRTNER, W. (1994). The primary structure of mantid opsin. *Gene* **143**, 227–231.

TOWNER, P., HARRIS, P., WOLSTENHOLME, A. J., HILL, C., WORM, K.

AND GÄRTNER, W. (1997). Primary structure of locust opsins: a speculative model which may account for ultraviolet wavelength light detection. *Vision Res.* **37**, 495–503.

WHITE, R. H., BANISTER, M. J. AND BENNETT, R. R. (1983*a*). Spectral sensitivity of screening pigment migration in the compound eye of *Manduca sexta*. *J. comp. Physiol.* A **l53**, 59–66.

WHITE, R. H., BROWN, P. K., HURLEY, A. K. AND BENNETT, R. R. (1983*b*). Rhodopsins, retinula cell ultrastructure and receptor potentials in the developing compound eye of the moth *Manduca sexta*. *J. comp. Physiol*. A **l50**, l53–l63.

WHITE, R. H., STEVENSON, R. D., BENNETT, R. R., CUTLER, D. E. AND HABER, W. A. (1994). Wavelength discrimination and the role of ultraviolet vision in the feeding behavior of hawkmoths. *Biotropica* **26**, 427–435.

XU, H. (1996). Expression of terminal fragments of cDNAs encoding *Manduca sexta* opsins in *Escherichia coli*. MS thesis, University of Massachusetts Boston, Boston, MA, USA.

YOKOYAMA, S. (1995). Amino acid replacements and wavelength absorption of visual pigments. *Molec. Biol. Evol*. **12**, 53–61.

YOKOYAMA, S. AND YOKOYAMA, R. (1996). Adaptive evolution of photoreceptors and visual pigments in vertebrates. *A. Rev. Ecol. Syst*. **27**, 543–567.