Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 6150–6154, July 1993
Cell Biology

Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus polyphemus*

W. CLAY SMITH*†‡, DAVID A. PRICE*, ROBERT M. GREENBERG*, AND BARBARA-ANNE BATTELLE*†

*Whitney Marine Laboratory and †Department of Neuroscience, University of Florida, Gainesville, FL 32086

Communicated by John E. Dowling, March, 19, 1993

ABSTRACT **cDNA clones encoding opsins from the lateral eyes and median ocelli of the horseshoe crab, *Limulus polyphemus*, were isolated from cDNA libraries. The opsin cDNAs obtained from the lateral eye and ocellar libraries code for deduced proteins with 376 amino acids. The two cDNAs are 96% identical at the nucleic acid level, differing primarily at the 3′ untranslated region, and are apparently the products of two separate genes. The deduced opsin proteins are 99% identical to each other, differing at only 5 amino acids. The opsins encoded by these cDNAs are most likely the protein moiety of the visible-wavelength rhodopsins in this animal. In the lateral eye, expression of the opsin gene is restricted to the photoreceptor cells of the ommatidia. Comparisons with opsins of other species show that the *Limulus* opsin proteins are most similar (53% identity) to the opsin from the R1–6 photoreceptors of flies.**

The role that rhodopsin plays in initiating the phototransduction cascade has been a model system for the function of guanine nucleotide-binding protein (G protein)-linked receptors. Visual pigments are fundamentally the same in all visual systems, being composed of a retinal chromophore covalently linked to a protein moiety with seven transmembrane segments (opsin; ref. 1). Consequently, the study of rhodopsin, and its use as a model, has been enhanced by the use of different organisms to study various aspects of visual pigments, taking advantage of each animal's unique physiology or morphology. For example, the large bovine eyes have provided rhodopsin in quantities sufficient for biochemistry (see ref. 2 for review), and *Drosophila* has provided a battery of mutants and transformants for genetic studies (e.g., ref. 3). Similarly, the horseshoe crab (*Limulus polyphemus*), with its large photoreceptors, which are easily accessible for electrophysiology, has provided insights into the activation of the transduction cascade by rhodopsin (4–6) and the inactivation of rhodopsin (7, 8).

Even though all visual pigments known in the animal kingdom have the same basic composition of opsin and a vitamin A-derived chromophore, there are differences in some characteristics of the visual pigments between species, such as spectral absorbance, kinetics, or interaction with subsequent enzymes in the cascade. Since the chromophore is almost universally the same, changes in the structure of the opsin moiety must predominantly account for the differences observed in the various rhodopsins. Analysis of the primary structure of visual pigments from different species continues to define and illuminate the amino acid domains that are involved in these differences. In this study, we present the amino acid sequences for opsin from the lateral eye and median ocelli of the horseshoe crab.§ *Limulus* is only the second class of arthropod from which the visual pigment has been sequenced. Further, to our knowledge this research is the first study of the primary sequence of any of the visual-system proteins in the horseshoe crab, despite the animal's longstanding significance in vision studies.

MATERIALS AND METHODS

The basic strategy for identifying and sequencing the cDNA for *Limulus* opsin was to generate cDNA libraries from *Limulus* lateral eyes and ocelli, amplify a portion of the opsin cDNA by polymerase chain reaction (PCR) using degenerate oligonucleotide primers, and obtain the full-length cDNA by screening the library for opsin clones by using the amplified cDNA as a homologous probe.

RNA Isolation. Lateral eyes from 15 horseshoe crabs, or median ocelli from 22 animals, were dissected from the animal and quick frozen in liquid nitrogen until needed. Total RNA was isolated by guanidinium thiocyanate extraction followed by centrifugation through a cushion of CsCl (9). The RNA pellet from this procedure was pelleted through a second cushion of CsCl to remove contaminating screening pigments. Poly(A)$^+$ RNA was prepared from the total RNA through selective binding to an oligo(dT)-cellulose column [Poly(A) Quik kit, Stratagene], with a yield of $\approx$0.5 μg per lateral eyes, and 0.05 μg per ocellus.

Library Construction. Separate cDNA libraries were constructed from poly(A)$^+$ RNA isolated from the lateral eyes and ocelli with the Superscript λ cDNA kit (GIBCO/BRL). The resultant lateral-eye library had 1.3×10^6 clones with an average insert size of 2 kb, and the ocellar library had 4.9×10^5 clones with an average insert size of 1 kb.

PCR. A degenerate oligonucleotide primer was designed against the amino acid sequence NPXXY (where X is a nonpolar amino acid) that is conserved in the seventh membrane-spanning segment of many G protein-linked receptors [5′-AA(T/C)(A/T)(C/G)II(C/G)CITCAA(T/C)CC(C/G)IT(C/G)IT(C/G)TA-3′, where I is deoxyinosine]. This primer, paired with a primer directed against the phage vector (5′-ACTCCTGGAGCCCGTCAGTATC-3′), was used to amplify the carboxyl terminus of the opsin cDNA by PCR. PCR products were subcloned into pCR (Invitrogen, San Diego) or pT7-Blue (Novagen, Madison, WI) plasmids and sequenced by the dideoxy termination method (10).

Library Screening. The cDNA libraries were screened for opsin clones by using a fragment of opsin cDNA that spanned helix 7 to the carboxyl terminus. The cDNA was labeled with [α-^{32}P]dCTP by random priming (11). Plaque lifts of the libraries were hybridized to the labeled probe in aqueous Denhardt's hybridization buffer (68°C) (12) and washed at high stringency (15 mM NaCl/1.5 mM sodium citrate/0.5% SDS, 68°C). Positive plaques were detected by autoradiography, and purified to homogeneity by replating. The cDNA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡To whom reprint requests should be sent at the present address: Department of Ophthalmology, Box 100284, J. Hillis Miller Health Center, Gainesville, FL 32610-0284.

§The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L03781 and L03782).

inserts were subcloned into pSPORT plasmid (GIBCO/BRL) for sequencing.

Northern Blot Analysis. Poly(A)$^+$ RNA (10 μg) from the lateral eyes was electrophoresed under denaturing conditions in a 1.1% agarose gel with the glyoxal method (13), transferred to nylon membrane, and crosslinked to the nylon by ultraviolet irradiation (120 mJ/cm^2). A cDNA probe [nt 730–1170 (see Fig. 1), labeled as above] was hybridized under the same conditions and stringency used to screen the library. Because of the size of the ocelli and the difficulty in obtaining tissue, a Northern blot of ocellar RNA was not performed for the ocellar opsin.

***In Situ* Hybridization.** Single-stranded cDNA was prepared for *in situ* hybridization by asymmetric PCR. Double-stranded cDNA between positions 1120 and 1241 (see Fig. 1) was first amplified for 35 cycles. Aliquots of the resulting double-stranded DNA were then cycled for an additional 25 cycles, using only one primer in each aliquot to generate sense and antisense single-stranded DNA (ssDNA). The ssDNA was gel purified (Qiaex; Qiagen, Studio City, CA), and the complementary strand was synthesized by random primed labeling in the presence of [α-[^{35}S]thio]dCTP. Unincorporated label was removed by column chromatography (Nuc-Trap, Stratagene). The labeled strand complementary to the mRNA was used as the probe, and the labeled strand with the same sense as the mRNA was used as a negative control.

Tissue was prepared for hybridization by fixing the lateral eyes for 30 min in 4% paraformaldehyde in phosphate-buffered saline (PBS: 137 mM NaCl/2.7 mM KCl/8.1 mM Na_2HPO_4/1.5 mM NaH_2PO_4) with 15% sucrose. Frozen sections (15 μm) were cut (Tissue-Tek OCT compound; Miles), fixed for 2 min with 4% paraformaldehyde in PBS, and dehydrated through an ethanol series. Hybridizations, washes, and emulsion autoradiography were conducted by the methods of Hafen and Levine (14). All photographs were taken with brightfield illumination due to the light-scattering effects of the pigment granules in the ommatidia.

RESULTS AND DISCUSSION

Isolation of cDNA Clones. A portion of the opsin cDNA was isolated from both the lateral-eye and ocellar libraries by PCR. The products from the lateral-eye cDNA library (278 bp) and from the ocellar cDNA library (304 bp) were sequenced to confirm their identity as opsin cDNA by comparison with known opsin sequences in the GenBank data base (15). This opsin cDNA was then used to screen the cDNA library to obtain full-length opsin clones (Fig. 1). At least two separate λ clones plus several PCR clones were sequenced on the forward and reverse strands to confirm the cDNA sequence. In total, 11 independent clones were sequenced for the lateral-eye opsin, and 12 clones for the ocellar opsin.

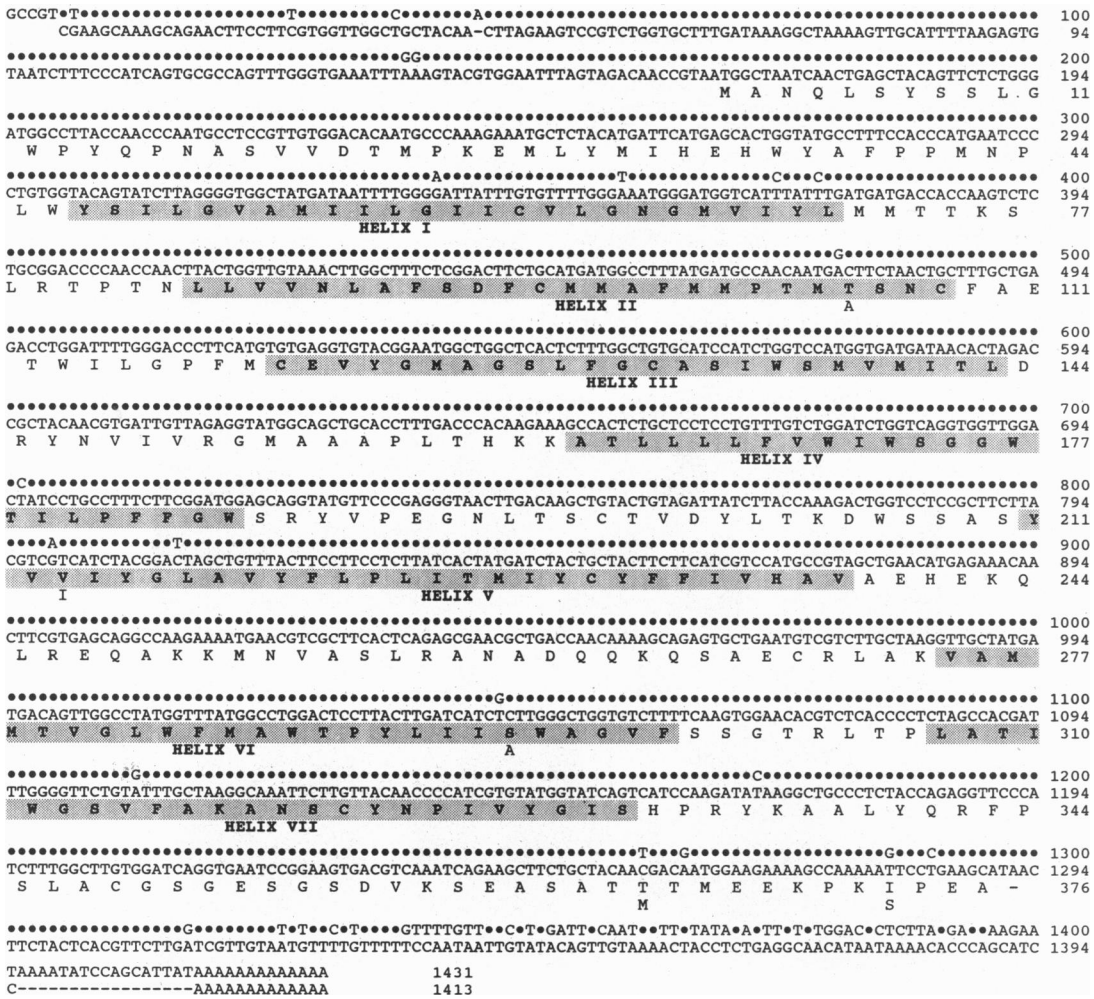

FIG. 1. cDNA sequences for the opsin from the lateral eyes (lower line) and median ocelli (upper line). Dots indicate sequence identity, and only where the two sequences differ is the nucleotide indicated. The deduced amino acid sequence for the lateral-eye opsin is listed below the cDNA sequences and the differences of the ocellar opsin are indicated below the sequence for the lateral eye opsin. Seven putative transmembrane segments are also indicated; the boundaries are based on alignments with bovine sequence, where the boundaries are more clearly established by additional studies.

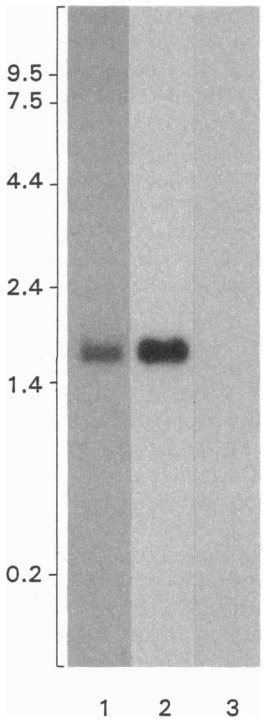

FIG. 2. Northern blot analysis of poly(A)$^+$ RNA (10 μg per lane) from the lateral eyes of *Limulus*. Lane 1, blot probed with opsin cDNA from the lateral eye, spanning helices V–VII; lane 2, blot probed with 3′ noncoding region of lateral-eye opsin cDNA; lane 3, blot probed with 3′ noncoding region of ocellar opsin cDNA. Molecular sizes of RNA markers are given in kilobases.

In the lateral-eye library, ≈2% of the plaque-forming units hybridized to the probe, consistent with the level of opsin message found in other visual systems (16). A lower level of hybridization was found in the ocellar library (≈0.1%). This lower level probably reflects two principal factors—a higher percentage of nonrecombinant phage in the ocellar library and a less pure tissue used for isolation of the ocellar RNA, since the ocelli are surrounded by a large number of pigment cells and connective tissue.

Northern blot analysis of poly(A)$^+$ RNA obtained from the lateral eyes of eight horseshoe crabs showed a single band corresponding to 1.5 kb (Fig. 2, lane 1).

Comparison of cDNA Clones. Two different nucleotide sequences were obtained for the opsin cDNA from the lateral eyes and median ocelli. The opsin cDNA from the lateral eye was 1413 bp long, encoding a deduced protein with 376 residues. This cDNA length corresponds well with the predicted size of 1.5 kb obtained from the Northern blot. The opsin cDNA from the median ocelli was 1431 bp long, also coding for a deduced protein with 376 amino acids. The two cDNA sequences are very similar (96% identity), differing primarily in the 3′ noncoding sequence.

The two opsin amino acid sequences deduced from the cDNA sequences are 99% identical, differing at only five amino acids. Polymorphisms at these sites were ruled out by sequencing a minimum of five clones for each site, using clones obtained from library screens and PCR. In all cases, the nucleotide differences at these positions were consistent. This consistency between clones also rules out the possibility of cloning artifacts. Of the five amino acid differences, three are conservative changes (Ala^{105} for Thr^{105} in helix II; Ile^{213} for Val^{213} in helix V; Ala^{294} for Ser^{294} in helix VI). The other two differences are at the carboxyl terminus (Met^{365} for Thr^{365} and Ser^{373} for Ile^{373}). In both cases, the proteins have a calculated molecular mass of 42 kDa.

Despite the high similarity of the cDNAs, the two opsin cDNAs are apparently encoded by different genes. Alternative splicing is ruled out by the few, but consistent, nucleotide differences found throughout the length of the two cDNA molecules. These differences are found in the 5′ noncoding region, as well as the coding region of the cDNAs. Consequently, while alternative splicing could account for the differences at the 3′ end, alternative splicing cannot account for the differences found throughout the rest of the cDNA molecule.

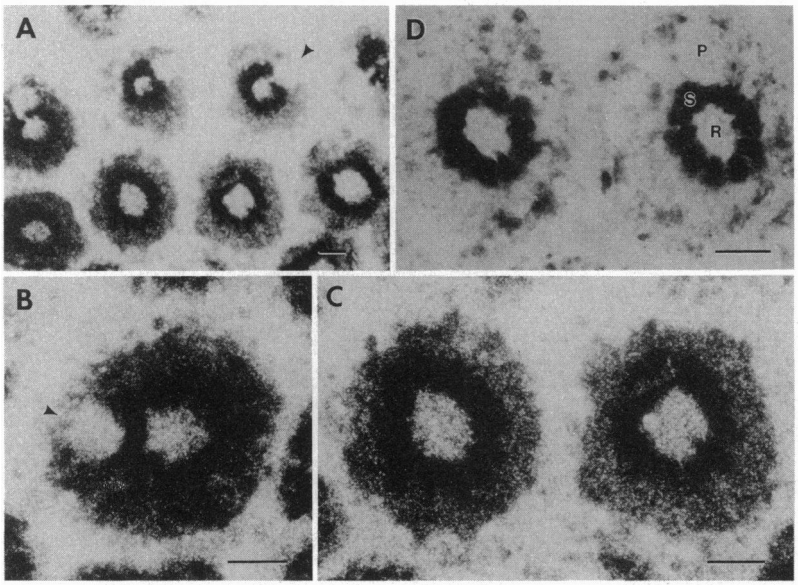

FIG. 3. Tissue localization of opsin mRNA in the lateral eyes of the horseshoe crab. Hybridization signals by the opsin-specific probe (^{35}S-labeled opsin cDNA) were detected by autoradiography. (*A*) Cross section through several ommatidia (arrowhead points to an eccentric cell body). (*B*) Higher magnification of cross section through a single ommatidium at the level where the eccentric cell body (arrowhead) is interposed between photoreceptor cells. (*C*) Cross section through two adjacent ommatidia, showing the hybridization signal only in the photoreceptor cell bodies of each ommatidium. (*D*) Control slide using ^{35}S-labeled probe with the same sense as the mRNA. P, photoreceptor cell body; S, screening pigments; R, rhabdom. (Bars = 50 μm.)

To demonstrate specific expression of each cDNA in its respective tissue, the 3′ noncoding region for each opsin cDNA was used to probe a Northern blot of poly(A)$^+$ RNA obtained from the lateral eye (Fig. 2, lanes 2 and 3). Only the cDNA for the lateral-eye opsin hybridized to the blot at the expected 1.5-kb size (lane 2). As additional evidence, these same probes were used to screen plaque lifts of both the lateral-eye and ocellar cDNA libraries. In this experiment, the lateral-eye cDNA probe hybridized only to plaques on the lift from the lateral-eye library, and the ocellar opsin cDNA probe hybridized only to the lift from the ocellar library. These data indicate that the lateral-eye opsin cDNA obtained from the cDNA library is specific to the lateral eye and that the ocellar opsin cDNA is specific to the median ocelli.

Both of the opsins described here are probably the protein moieties of visible-wavelength rhodopsins for the following reasons. First, only one visual pigment (λ_{max} of 520 nm) has been identified in the lateral eyes of *Limulus* (17). During the course of this set of experiments, only one opsin cDNA sequence was observed even though portions of 11 different cDNA clones were sequenced. Further, as discussed in the following section, the lateral-eye opsin has a potential Schiff base counterion, which would shift the absorption spectrum of the visual pigment from the ultraviolet to the visible range. By the same logic, the ocellar opsin cDNA also probably encodes a visible-wavelength rhodopsin, since the deduced amino acid sequence differs from the lateral-eye opsin only in residues that are noncharged in the transmembrane domains, and consequently should not significantly affect the wavelength of maximum absorption. Although at least two classes of visual pigments exist in the ocelli (λ_{max} of 530 nm and 360 nm; ref. 18), only the one opsin was isolated from the ocellar library.

Tissue Distribution. Opsin message for the lateral eye was localized by *in situ* autoradiography using cDNA coding for the protein spanning a portion of helix VII and the carboxyl tail. The message localized exclusively to the photoreceptor cells of each ommatidium (Fig. 3). The eccentric cell, which has its cell body intercalated between two photoreceptor cells

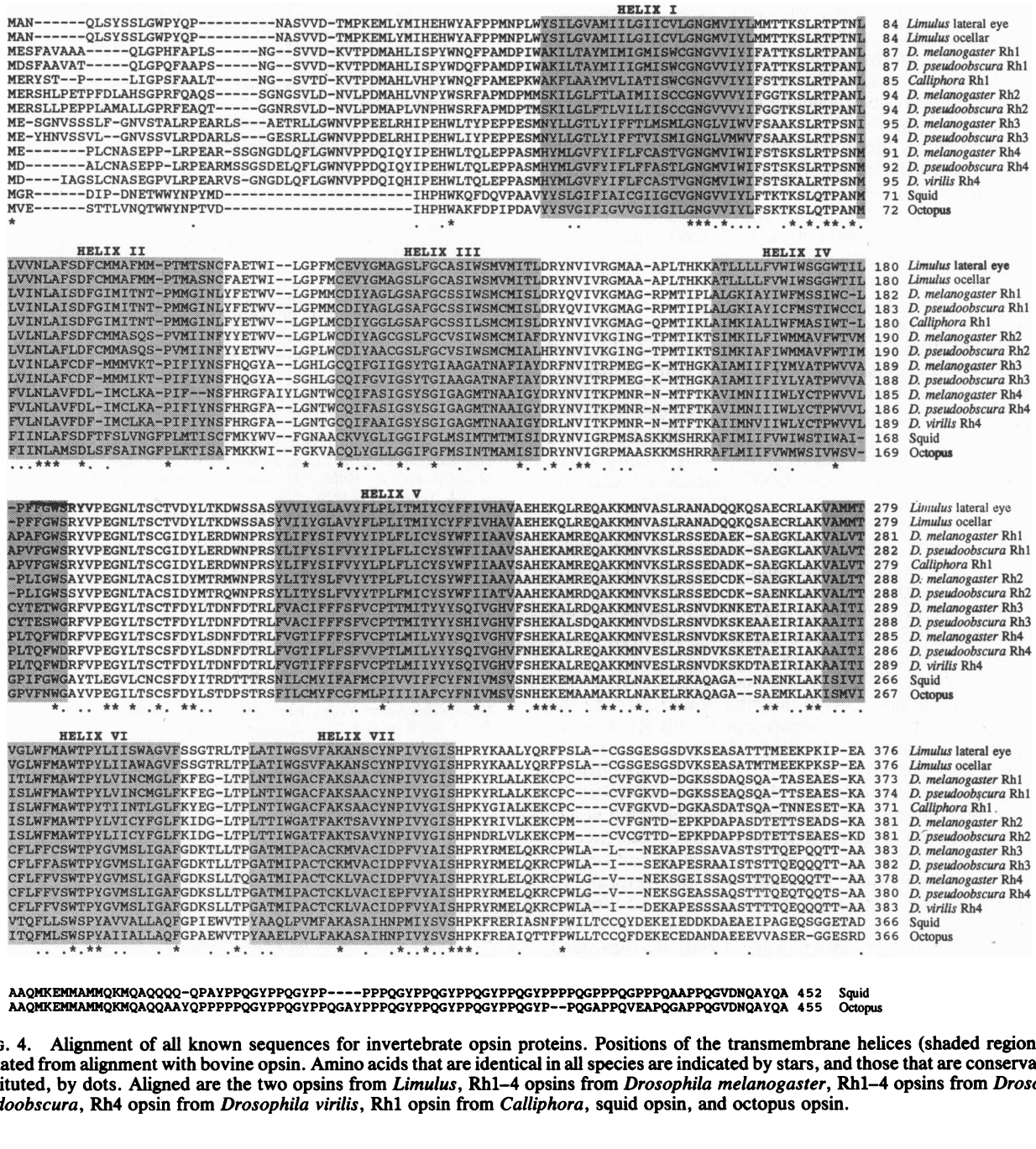

FIG. 4. Alignment of all known sequences for invertebrate opsin proteins. Positions of the transmembrane helices (shaded regions) are estimated from alignment with bovine opsin. Amino acids that are identical in all species are indicated by stars, and those that are conservatively substituted, by dots. Aligned are the two opsins from *Limulus*, Rh1–4 opsins from *Drosophila melanogaster*, Rh1–4 opsins from *Drosophila pseudoobscura*, Rh4 opsin from *Drosophila virilis*, Rh1 opsin from *Calliphora*, squid opsin, and octopus opsin.

in each ommatidium, did not stain for the opsin message (Fig. 3 *A* and *B*). The eccentric cell is a nonphotosensitive cell that is electrically coupled to the photoreceptors and generates action potentials in response to light (19). The dark central ring in all the ommatidia (Fig. 3) consists of the screening pigment granules of the light-adapted eye condensed around the central rhabdom (not preserved by this fixation method).

Species Comparisons. When aligned with all other opsin protein sequences in the GenBank data base (September 1992) (method of refs. 20 and 21), the *Limulus* opsin sequences are most similar to the Rh1 and Rh2 opsins of flies. Both the lateral eye and ocellar opsins share a 53% identity with Rh1 and 49% identity with Rh2 from *Drosophila melanogaster*. Because the horseshoe crab represents only the second class of arthropod from which opsin has been sequenced, it is appropriate to consider a comparison of invertebrate opsins in more detail. Alignment of the 14 known invertebrate sequences [10 from flies (3, 22–28), 2 from cephalopods (29, 30), and 2 from *Limulus*] shows conservation of several domains (Fig. 4). The highest degree of conservation is in the cytoplasmic loops, especially the loops between helices I and II and between helices V and VI. Loops 3–4 and 5–6 are known to interact with the G protein in vertebrate rhodopsins (31, 32). Interestingly, there is also a strong conservation of extracellular loop 4–5, for which there is no known function.

This alignment of the invertebrate opsins also reveals in *Limulus* opsins conservation of some of the amino acids that have significant functional roles in other species: Asn-Ala-Ser[19], a potential glycosylation site (33); Cys[120]/Cys[197], which form a disulfide bond (34); Lys[318], the retinal binding site (35); and nine serine and threonine residues in the carboxyl terminus, potential sites for phosphorylation by rhodopsin kinase (36).

In vertebrate opsins a glutamic acid in helix III (Glu[113] of bovine rhodopsin) is the Schiff base counterion (37, 38). Both *Limulus* proteins also have a glutamic acid (Glu[121]) near the location of this glutamic acid. However, in *Limulus* this glutamic acid is adjacent to one of the conserved cysteines (Cys[120]) that probably participates in a disulfide bond with another cysteine residue (Cys[197]). Cys[197] is clearly in an extracellular loop. This places Glu[121] outside or near the boundary of helix III; it is thus unlikely that Glu[121] is the Schiff base counterion in *Limulus*. A possible alternative counterion in helix III is Tyr[123]. According to the model proposed by Hall *et al.* (29), an induced dipole is created in the polar tyrosine, thus delocalizing the Schiff base charge. This tyrosine is conserved in all invertebrate opsins, except the opsins in the ultraviolet visual pigments from flies, where a Schiff base counterion is not expected.

The close similarity of the ocellar and lateral-eye opsins in *Limulus* is quite different from the situation in flies. In the fruit fly, Rh1 (the opsin for photoreceptor cells R1–6 in the compound eye) is only 69% identical to Rh2 (the opsin for the ocelli). Further, although both fly rhodopsins are visible-wavelength visual pigments, the rhodopsin in R1–6 has an absorption maximum at 480 nm and that in the ocelli at 425 nm (39). Thus, the difference in amino acid sequence is reflected by a difference in absorption spectra. In contrast, in *Limulus* the absorption maximum of the lateral-eye rhodopsin (520 nm; ref. 17) differs from the absorption maximum of the visible-wavelength ocellar opsin (530 nm; ref. 18) by only 10 nm, in agreement with the observed similarity of the ocellar and lateral-eye opsins.

An analysis of phylogenetic relationships based on the percentage of amino acid identity between the currently known opsin sequences (average linkage cluster analysis; ref. 40) shows that the *Limulus* visual pigments cluster, expectedly, with the other arthropod opsins among the invertebrates. The high identity of the *Limulus* lateral-eye and ocellar opsins suggests a possible gene duplication event. If this is the case, then the duplication must have occurred recently compared with other known opsin duplication events. A logical comparison is to the duplication leading to the red and green pigments of human cones, which have 98% identity (41). In this case, the duplication is estimated to have occurred 40 million years ago, suggesting a similar age for the possible *Limulus* duplication.

This research was supported by grants from the National Institutes of Health (EY06454 to W.C.S.), National Science Foundation (BNS-8909052 to B.-A.B. and DIR-8819005 to the Whitney Lab), and University of Florida Division of Sponsored Research (Research Development Award to D.A.P.). The DNA Synthesis Core (University of Florida Interdisciplinary Center for Biotechnology Research) synthesized the oligonucleotide primers used in this research.

1. Hubbard, R. & Wald, G. (1952) *J. Gen. Physiol.* **36**, 269–270.
2. Hargrave, P. A. (1982) *Prog. Retinal Res.* **1**, 1–51.
3. O'Tousa, J. E., Baehr, W., Martin, R. L., Hirsh, J., Pak, W. L. & Applebury, M. L. (1985) *Cell* **40**, 839–850.
4. Brown, J., Rubin, L., Ghalayini, A., Tarver, A., Irvine, R., Berridge, M. & Anderson, R. (1984) *Nature (London)* **311**, 160–163.
5. Fein, A. (1986) *Trends NeuroSci.* **9**, 110–114.
6. Kirkwood, A., Wiener, D. & Lisman, J. E. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3872–3876.
7. Levine, E., Crain, E. & Robinson, P. (1987) *J. Gen. Physiol.* **90**, 575–586.
8. Richard, E. A. & Lisman, J. E. (1992) *Nature (London)* **356**, 336–338.
9. MacDonald, R. J., Swift, G. H., Przybyla, A. E. & Chirgwin, J. M. (1987) *Methods Enzymol.* **152**, 219–227.
10. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
11. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
12. Church, G. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
13. McMaster, G. K. & Carmichael, G. G. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4835–4838.
14. Hafen, E. & Levine, M. (1986) in *Drosophila: A Practical Approach*, ed. Roberts, D. B. (IRL, Oxford), pp. 139–157.
15. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
16. Archer, S. N., Lythgoe, J. N. & Hall, L. (1992) *Proc. R. Soc. Lond. B* **248**, 19–25.
17. Hubbard, R. & Wald, G. (1960) *Nature (London)* **186**, 212–215.
18. Nolte, J. & Brown, J. E. (1972) *J. Gen. Physiol.* **59**, 186–200.
19. Behrens, M. E. & Wulff, V. J. (1965) *J. Gen. Physiol.* **48**, 1081–1093.
20. Higgins, D. G. & Sharp, P. M. (1988) *Gene* **73**, 237–244.
21. Higgins, D. G. & Sharp, P. M. (1989) *Comp. Appl. Biosci.* **5**, 151–153.
22. Zuker, C. S., Montell, C., Jones, K., Laverty, T. & Rubin, G. M. (1987) *J. Neurosci.* **7**, 1550–1557.
23. Montell, C., Jones, K., Zuker, C. & Rubin, G. M. (1987) *J. Neurosci.* **7**, 1558–1566.
24. Zuker, C. S., Cowman, A. F. & Rubin, G. M. (1985) *Cell* **40**, 851–858.
25. Cowman, A. F., Zuker, C. S. & Rubin, G. M. (1986) *Cell* **44**, 705–710.
26. Carulli, J. P. & Hartl, D. L. (1992) *Genetics* **132**, 193–204.
27. Neufeld, T. P., Carthew, R. W. & Rubin, G. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10203–10207.
28. Huber, A., Smith, D. P., Zuker, C. S. & Paulsen, R. (1990) *J. Biol. Chem.* **265**, 17906–17910.
29. Hall, M. D., Hoon, M. A., Ryba, N. J. P., Pottinger, J. D. D., Keen, J. N., Saibil, H. R. & Findlay, J. B. C. (1991) *Biochem. J.* **274**, 35–40.
30. Ovchinnikov, Y. A., Abdulaev, N. G., Zolotarev, A. S., Artomonov, I. D., Bespalov, I. A., Dergachev, A. E. & Tsuda, M. (1988) *FEBS Lett.* **232**, 69–72.
31. Franke, R. R., Konig, B., Sakmar, T. P., Khorana, H. G. & Hofmann, K. P. (1990) *Science* **250**, 123–125.
32. Franke, R. R., Sakmar, T. P., Graham, R. M. & Khorana, H. G. (1992) *J. Biol. Chem.* **267**, 14767–14774.
33. Struck, D. K. & Lennarz, W. J. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans*, ed. Lennarz, W. J. (Plenum, NY), pp. 35–83.
34. Karnik, S. S., Sakmar, T. P. & Chen, H.-B. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8459–8463.
35. Bownds, D. (1967) *Nature (London)* **216**, 1178–1181.
36. Wilden, U. & Kuhn, H. (1982) *Biochemistry* **21**, 3014–3022.
37. Sakmar, T. P., Franke, R. R. & Khorana, H. G. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8309–8313.
38. Nathans, J. (1990) *Biochemistry* **29**, 9746–9752.
39. Stark, W. S., Frayer, K. L. & Johnson, M. A. (1979) *Biophys. Struct. Mech.* **5**, 197–209.
40. Sneath, P. H. A. & Sokal, R. R. (1973) *Numerical Taxonomy* (Freeman, NY).
41. Nathans, J. (1989) *Sci. Am.* **260**, 42–49.