Volume 315, number 3, 287–292 FEBS 11985 January 1993

Primary structure of crayfish visual pigment deduced from cDNA

Takahiko Hariyama[a], Koichi Ozaki[b], Fumio Tokunaga[b] and Yasuo Tsukahara[a,c]

[a]*Research Center for Applied Information Sciences, Tohoku University, Katahira 2, Aoba-ku, Sendai 980, Japan,* [b]*Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan* and [c]*Photodynamics Research Center, The Institute of Physical and Chemical Research (Riken), Aoba-dai, Aoba-ku, Sendai 980, Japan*

Received 15 October 1992; revised version received 19 November 1992

The primary structure of opsin of the crayfish *Procambarus clarkii* has been deduced from the cDNA sequence. The opsin is composed of 376 amino acid residues including all the conservative residues characteristic of other members of the rhodopsin family. Comparison of sequences of all known opsins reveals that the major *Drosophila* rhodopsin is more similar to the crayfish rhodopsin than to the *Drosophila* UV-sensitive pigments. The phylogenetic trees of invertebrate opsins are constructed.

Rhodopsin; Opsin; Photoreceptor cell; Retina; UV-receptor; *Drosophila*; Crayfish (*Procambarus clarkii*)

1. INTRODUCTION

Since the isolation and sequencing of genomic and complementary DNA clones encoding the apoproteins of human blue-, green- and red-sensitive cone visual pigments [1], studies on the molecular basis of vertebrate color vision have developed rapidly. Deduced amino acid sequences of cone visual pigments, as well as those of rod pigments, have been reported in a chicken [2,3], a nocturnal lizard [4] and fishes [5,6], and studies on the molecular evolution of these pigments suggest that the vertebrates acquired color vision before scotopic vision based on rhodopsin [3].

In invertebrates, studies on color discrimination have been carried out mainly on insects [7], revealing that they have a well-developed multichromatic visual system (color vision) in the range between near-UV and red. There have been many studies to date on the Crustacea, too [8,9]. The crayfish, *Procambarus clarkii*, is a well-known experimental animal with a pair of well developed compound eyes. Each ommatidium has 8 photoreceptor cells, seven of which (R1–7) form the main rhabdom with maximum spectral sensitivity at about 600 nm [10]. The eighth photoreceptor cell (R8) is located distal to the main rhabdom and absorbs blue light (λ_{max} = 440 nm) [11]. On the other hand, the eye of crayfish possesses two kinds of chromophores, retinal (A1) and 3-dehydroretinal (A2) [12], and there is a seasonal variation of the A2 content [13], causing changes in the spectral sensitivity of the photoreceptor cells [14]. Recently, both A1- and A2-pigments have been characterized spectrally, and shown to have absorbance peaks at 533 nm and 567 nm, respectively [15]. These results therefore suggest that the crayfish possibly has a multichromatic visual system in the range between blue and red, but possesses no UV-receptors. In other species of Crustacea, however, some evidences suggesting the existence of a UV receptor have also been reported [16].

In the present study, we have determined the sequence of the crayfish opsin cDNA to deduce the primary structure of the protein. By comparing this sequence with those of other opsins, it is suggested that the UV reception in the arthropod had already been established in a common ancestor of Crustacea and Insecta before these two classes diverged.

Correspondence address: Koichi Ozaki, Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan. Fax: (81) (6) 845 7966.

2. MATERIALS AND METHODS

Six hundred crayfishes, *Procambarus clarkii*, were captured locally, and their retinas were surgically separated from the dioptric apparatus and the optic ganglia. The retinas were immediately frozen in liquid nitrogen, and about 200 μg of total RNA was extracted using the guanidinium thiocyanate method [17]. Poly(A)$^+$ RNA was isolated from total RNA by the use of Oligotex-dT30, and used for cDNA synthesis.

Four cDNA fragments which partially overlap each other were amplified by polymerase chain reaction (PCR), and were reconstructed to give the complete sequence of the cDNA from a single species of opsin mRNA (Fig. 1). First, three kinds of single-stranded cDNAs (ss-cDNAs) were made as follows by using the primers IVRH-R8 and CFRH-R4, and the dT-tailed pUC9 vector-primer. The reaction mixture was composed of 100 ng of denatured mRNA, 100 ng of oligonucleotide primer or vector-primer, 30 units of AMV reverse transcriptase, 50 mM Tris-HCl (pH 8.3), 100 mM KCl, 10 mM $MgCl_2$, 4 mM dithiothreitol, 40 units of RNasin and 1.5 mM each dNTP in a total volume of 30 μl, and was incubated at 42°C for 2 h. PCR was first carried out between IVRH-F5 and IVRH-R8 primers. For template DNA, the ss-cDNA was made with the IVRH-R8 primer, as

described above. The second and third PCRs were performed between CFRH-F1 and M13-RV, and between IVRH-F1 and CFRH-R1 by using the ss-cDNA synthesized on the dT-tailed pUC9 vector-primer. The former reaction amplified the cDNA fragment corresponding to the 3'-end of opsin mRNA. In order to amplify the 5'-end cDNA, double-stranded cDNA (ds-cDNA) fragments were synthesized from the ss-cDNA which was made with the CFRH-R4 primer, and inserted into the *Sma*I site of pUC118 plasmid. PCR was then carried out between the primers of M13-M4 and CFRH-R3. For each PCR, the mixture contained 100 pmol of each primer, 20 nmol of each dNTP, 4 units Tth DNA polymerase, 10 mM Tris-HCl (pH 8.9), 1.5 mM $MgCl_2$, 80 mM KCl, 50 μg BSA, 0.1% sodium cholate, 0.1% Triton X-100 and sterilized water, added to 5 μl of the template ss- or ds-cDNA solution to make a final volume of 100 μl. To amplify DNA, 30 cycles of denaturation, annealing and extension were completed. The annealing reaction was performed at 55°C for 1.5 min, followed by extension at 72°C for 2–3.5 min and denaturation at 94°C for 1 min.

The products of PCRs were recovered by ethanol precipitation, treated with restriction enzymes (*Eco*RI and *Bam*HI) which cut the ends of primers, and separated by agarose gel electrophoresis. cDNA fragments which formed major bands on the agarose gels were then isolated and cloned into pGEM7zf(+) for sequencing. At least 5 clones were isolated and sequenced for each cDNA band.

ODEN programs of the DNA Research Center, National Institute of Genetics, Japan were used for a multiple alignment of amino acid sequences of opsins, and for the construction of phylogenetic trees of the sequences. For alignment, N- and C-terminal regions of opsins presumed to lie outside the lipid bilayer were omitted, because these regions of opsin are too variable between animals in different phyla. Phylogenetic trees were constructed by the unweighted pair-group method with arithmetic mean (UPGMA) [18], the neighbor-joining method (NJ) [19] and the maximum parsimony method [20].

3. RESULTS AND DISCUSSION

Using 8 oligonucleotide primers (Fig. 1a), we amplified 4 partially overlapping cDNA fragments as shown in Fig. 1b. We concluded that these cDNA fragments arose from mRNA of a single kind of opsin because of

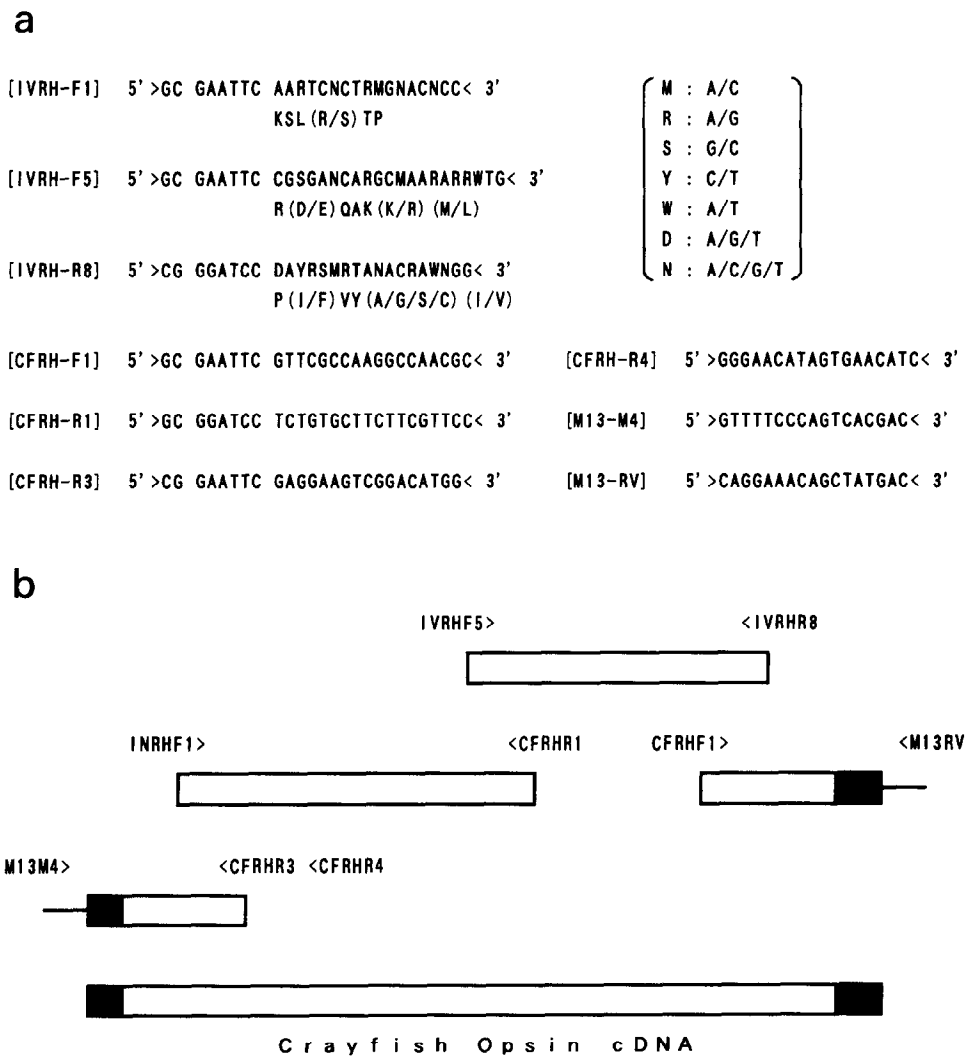

Fig. 1. (a) Synthetic oligonucleotide primers for crayfish opsin cDNA synthesis and PCR. IVRH primers correspond to amino acid sequences highly conserved in invertebrate opsins. Each CFRH primer has a unique sequence in the crayfish opsin cDNA. M13-M4 and -RV are the universal primers for pUC/M13. All IVRH and CFRH primers (except CFRH-R4) have an *Eco*RI or *Bam*HI recognition sequence. (b) Cloning strategy for the crayfish opsin cDNA. The complete cDNA was reconstituted from four fragments partially overlapping each other. Filled bars represent noncoding regions.

the following reasons. (i) The nucleotide sequences of overlapping regions (minimum 10 nucleotides except the primers, Fig. 2) were in complete agreement, although the regions are not conservative in any combinations of other known opsins. (ii) Except for the first PCR, between IVRH-F5 and IVRH-R8, we used the oligonucleotides with cDNA-specific sequence as one of PCR primers, and performed PCRs at high annealing temperature (55°C) to avoid nonspecific annealing. To determine the sequences of the 5′- and 3′-ends, we constructed the cDNA in pUC plasmids, and carried out PCRs by using pCU/M13 universal primers. This method gave a clear band of amplified cDNA on an agarose gel, and the cDNA fragments were easily cloned for sequencing. We also tried the RACE protocol [21] for amplification of the 5′-end, but obtained poorly defined bands which could not be cloned.

In the first PCR, we isolated 8 cDNA clones, all of

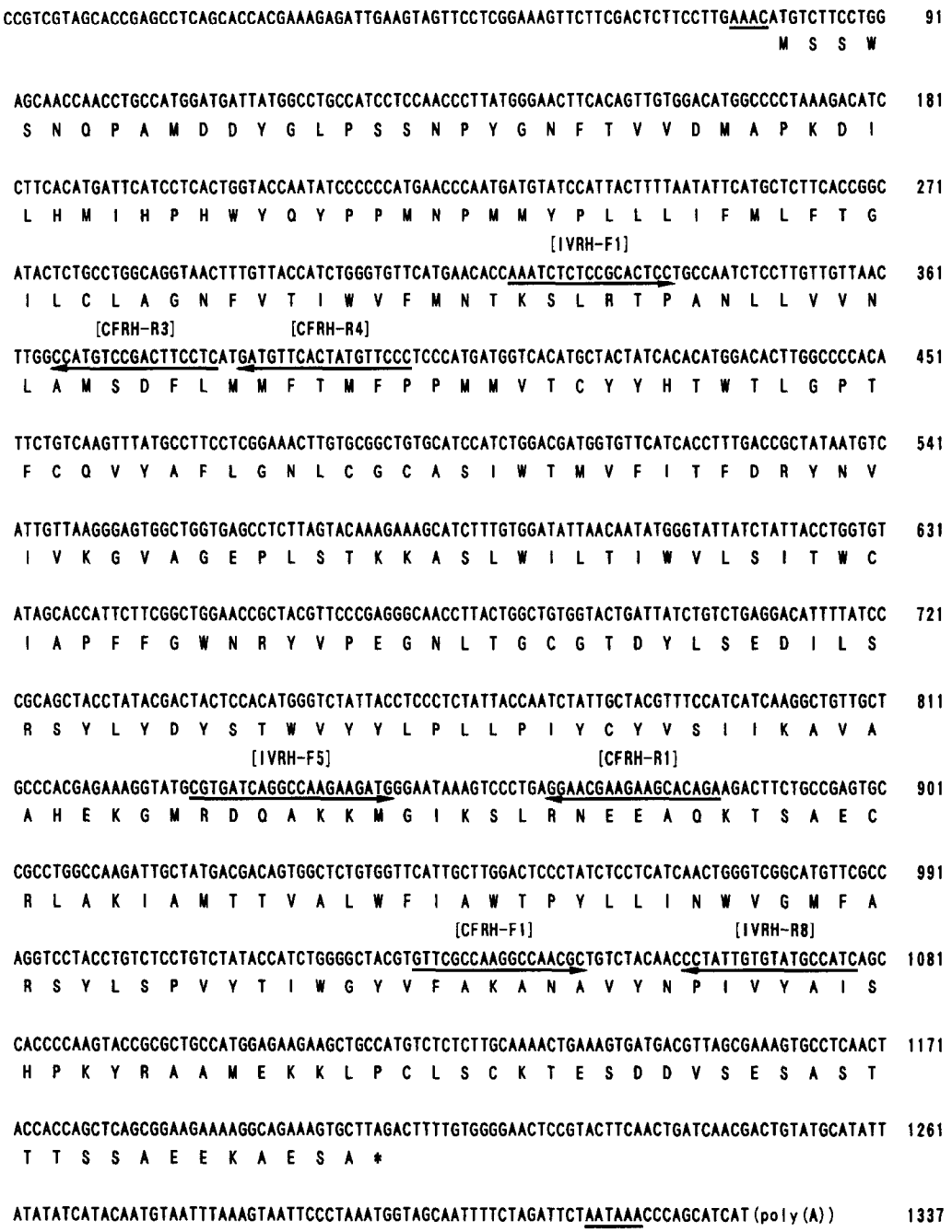

Fig. 2. cDNA nucleotide and deduced amino acid sequences of crayfish opsin. Consensus sequences for the translation initiation (AAAC) and the polyadenylation (AATAAA) signals are underlined. Horizontal arrows indicate the positions of primers.

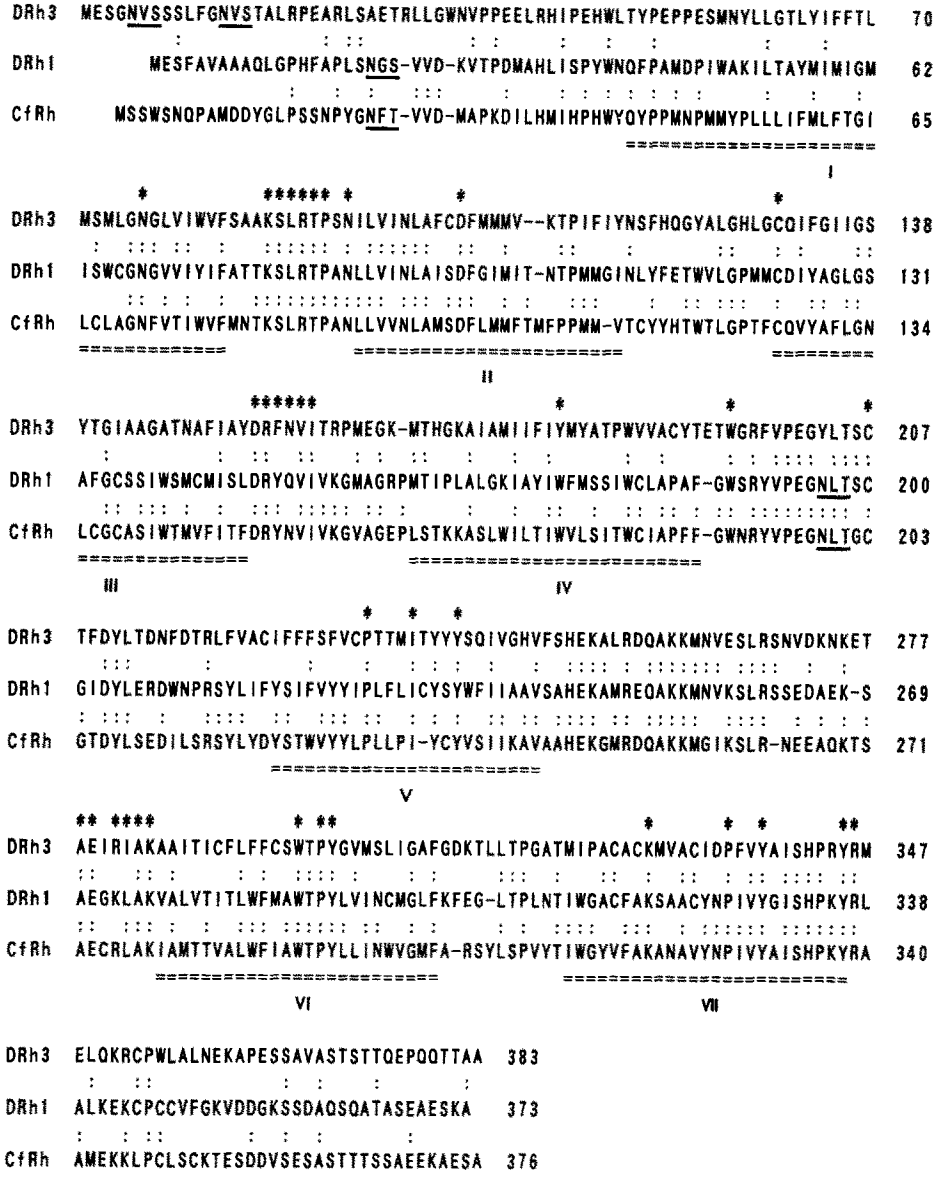

Fig. 3. Comparison of amino acid sequences of *Drosophila* major rhodopsin (DRh1) [22] with crayfish rhodopsin (CfRh) and *Drosophila* UV-sensitive pigment (DRh3) [24]. (:) indicates identical amino acid residues. (*) indicates the amino acid residues conserved in all members of the opsin family (see text). The predicted transmembrane segments and glycosylation sites are underlined.

which encoded the identical opsin polypeptide. Because an ommatidium of the crayfish compound eye contains 7 major and 1 blue-sensitive photoreceptors, it is strongly suggested that the cloned cDNA arise from opsin mRNA of the major photoreceptor cells. However, we cannot completely exclude the possibility that the cDNA encodes the blue-sensitive pigment, until the protein is functionally characterized.

The crayfish opsin cDNA consists of 1,337 nucleotides and a poly(A) tail. The polyadenylation signal (AATAAA; 1,321–1,326) was found at 17 bases upstream from the start of the poly(A) tail (Fig. 2). The cDNA sequence revealed a single open reading frame of 1,128 bp starting from the first ATG (81–83) and encoding 376 amino acids (Fig. 2). Because this first ATG has the preceding nucleotides (AAAC; 77–80) matching the consensus sequence for the *Drosophila* translation initiation sequence [22], we chose the ATG as the translational start codon. Based on the predicted sequence, the molecular weight of crayfish opsin was calculated to be 42,761. This value is rather larger than that estimated by SDS-PAGE (37,000), but such a discrepancy has been reported widely for other membrane proteins [23–25].

Like other known opsins, crayfish opsin is predicted to have seven trans-membrane domains (by the analysis

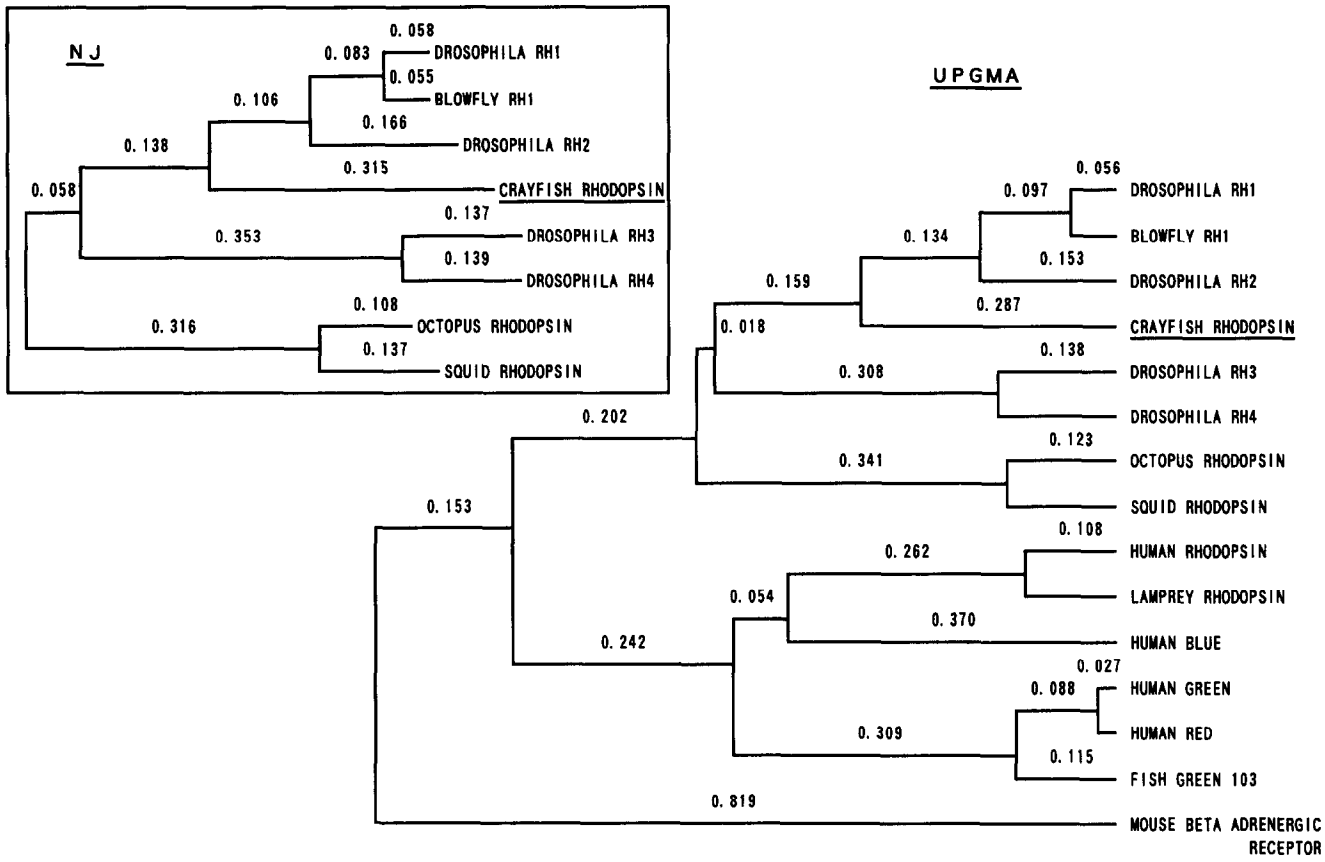

Fig. 4. Phylogenetic trees of visual pigments with branch lengths of evolutionary distance. In the NJ method (inset), the branch length near the deepest root was determined by a tree including vertebrate rhodopsins (not shown). Both the UPGMA and NJ methods give a similar branch point for *Drosophila* UV pigments (Rh3 and Rh4) from other arthropod rhodopsins.

of hydropathicity; data not shown), and possesses the appropriate conserved amino acid residues (Fig. 3). Lys-321 is likely to be the site of attachment of the chromophore, retinal. The two cysteine residues which have been proposed to form a structurally important disulphide bond in other opsins have equivalents (Cys-126 and Cys-203) in crayfish opsin. Series of residues in the cytoplasmic loops are thought to be responsible for binding and activation of G-protein, and are conserved in all opsins: in crayfish opsin, these residues are found in the corresponding positions, Lys-82–Asn-89, Asp-150–Ile-155, and Ala-272–Lys-278. In addition, some residues which seem to be important for protein structure (Asn-71, Asp-99, Trp-177, Trp-191, Pro-229, Ile-233, Tyr-236, Trp-291, Pro-293, Tyr-294, Pro-328, Tyr-331, Tyr-338, Arg-339) are conserved. Two possible *N*-glycosylation sites are located in the N-terminal region (Asn-23), and in the loop between helices 4–5 (Asn-199). Serine and threonine residues as potential phosphorylation sites are found in the C-terminal region (12 residues) and in the loop between helices 5–6 (5 residues).

Crayfish opsin is most similar to the opsin (Rh1) from photoreceptors R1–6 of *Drosophila* and *Calliphora*, with 53% amino acids identical. Surprisingly, comparison of the amino acid sequences of all known opsins reveals that the major *Drosophila* rhodopsin (Rh1) is more similar to the crayfish rhodopsin than to the *Drosophila* UV-sensitive pigments (Rh3 and Rh4). We constructed phylogenetic trees of opsins by means of three different methods. Since the result of the maximum parsimony method closely resembles to that of the NJ method, phylogenetic trees by UPGMA and NJ methods are shown in Fig. 4. In both expected-distance (UPGMA) and realized-distance (NJ) trees, it is shown that the *Drosophila* UV pigments diverged from other arthropod visual pigments significantly earlier than crayfish rhodopsin. This suggests that the UV reception in the arthropod had already been established in a common ancestor of Crustacea and Insecta before these two classes diverged. Recently, it has been shown that the vertebrates acquired color vision before scotopic vision [3]. Therefore, it seems likely that animals generally acquired the ability to detect various wavelengths of environmental light before evolving the ability to function at various intensities of light.

Acknowledgements. We thank Dr. L. Liao and other members of our laboratories for their assistance of capturing crayfishes and isolating retinas. We thank also Dr. I.G. Gleadall for his comments on the manuscript. This work was supported in part by Grant-in-Aid for General Scientific Research to T.H. and Grants-in-Aid for Scientific Research on Priority Areas to Y.T. and K.O. from The Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Nathans, J., Thomas, D. and Hogness, D.S. (1986) Science 232, 193–202.
[2] Kuwata, O., Imamoto, Y., Okano, T., Kokame, K., Kojima, D., Matsumoto, H., Morodome, A., Fukada, Y., Shichida, Y., Yasuda, K., Shimura, Y. and Yoshizawa, T. (1990) FEBS Lett. 272, 128–132.
[3] Okano, T., Kojima, D., Fukada, Y., Shichida, Y. and Yoshizawa, T. (1992) Proc. Natl. Acad. Sci. USA 89, 5932–5936.
[4] Kojima, D., Okano, T., Fukada, Y., Shichida, Y., Yoshizawa, T. and Ebrey, T.G. (1992) Proc. Natl. Acad. Sci. USA 89, 6841 6845.
[5] Yokoyama, R. and Yokoyama, S. (1990) Vision Res. 30, 807–816.
[6] Yokoyama, R. and Yokoyama, S. (1990) Proc. Natl. Acad. Sci. USA 87, 9315–9318.
[7] Peitsch, D., Fietz, A., Hertel, H., De Souza, J., Ventura, D.F. and Nenzel, R. (1992) J. Comp. Physiol. A 170, 23–40.
[8] Menzel, R., in: Handbook of Sensory Physiology (H. Autrum, Ed.) vol. VII/6A, Springer, Berlin, 1979, pp. 503–588.
[9] Hariyama, T., Meyer-Rochow, V.B. and Eguchi, E. (1986) J. Exp. Biol. 123, 1–26.
[10] Nosaki, H. (1969) Z. Vergl. Physiologie 64, 318–323.
[11] Cummins, D. and Goldsmith, T.H. (1981) J Comp. Physiol 142, 199–202.
[12] Suzuki, T., Makino-Tasaka, M. and Eguchi, E. (1984) Vision Res. 24, 783–787.
[13] Suzuki, T., Arikawa, K. and Eguchi, E (1985) Zool. Sci. 2, 455–461.
[14] Hariyama, T. and Tsukahara, Y (1988) Comp. Biochem. Physiol. 91A, 529–533.
[15] Zeiger, J. and Goldsmith, T.H. (1989) Vision Res. 29, 519–527.
[16] Cronin, T.W. (1989) J. Comp. Physiol. A 164, 737–749.
[17] Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M. (1977) Science 196, 1313–1319.
[18] Sokal, R.R. and Michener, C.D. (1958) Univ. Kansas Sci Bull. 28, 1409–1438.
[19] Saitou, N. and Nei, M. (1987) Mol Biol. Evol. 4, 406–425.
[20] Dayhoff, M.O and Park, C.M., in: Atlas of Protein Sequence and Structure (M.O. Dayhoff, Ed.), vol 4, Natl. Biomed. Res Found., Silver Springs, MD, 1969, pp. 7–16.
[21] Frohman, M A., Dush, M.K. and Martin, G.R. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[22] Cavener, D.R. (1987) Nucleic Acids Res. 15, 1353–1361.
[23] O'Tousa, J.E., Baehr, W., Martin, R.L., Hirsh, J., Pak, W L. and Applebury, M.L. (1985) Cell 40, 839–850.
[24] Hara-Nishimura, I., Matsumoto, T., Mori, H., Nishimura, M., Hara, R. and Hara, T. (1990) FEBS Lett. 271, 106–110
[25] Zuker, C.S., Montell, C., Jones, K., Laverty, T. and Rubin, G.M. (1987) J. Neurosci. 7, 1550–1557.