A novel and ancient vertebrate opsin

Bobby G. Soni, Russell G. Foster*

Imperial College of Science, Technology and Medicine, Department of Biology, Prince Consort Road, London SW7 2BB, UK

Received 5 February 1997; revised version received 25 February 1997

Abstract We describe the identification of a novel opsin gene isolated from the eyes of Atlantic salmon. The cDNA sequence predicts a protein that has the key features of an opsin, but shows only 32–42% amino acid identity to the known opsin families. Phylogenetic analysis suggests that this opsin is a member of a hitherto unrecognised opsin family that diverged early in the evolution of vertebrate photopigments. We have tentatively called this opsin family the vertebrate ancient (VA) opsins. The identification of VA opsin may ultimately help to resolve some of the uncharacterised photoreceptor functions of the eye, which include the regulation of circadian rhythms, pupil size and corneal pigmentation.

Key words: Opsin evolution; Vertebrate photopigment; Cone photopigment; Rod photopigment; Teleost fish

1. Introduction

All animal photopigments consist of a membrane-embedded protein (opsin) connected by a Schiff base linkage to a light-sensitive chromophore (11-*cis*-retinoid) [1]. During the course of vertebrate evolution, amino acid changes (along with gene duplications) have led to the formation of several distinct opsin families. Opsins from different families show approximately 45% amino acid identity and 65% amino acid similarity, whilst members within an opsin family show a much higher level of identity (85%) and similarity (95%), even from species as distant as goldfish and human.

Okano et al. [2] have proposed a phylogenetic tree in which the long wavelength (group L) and short wavelength (group S) opsins diverged from a common ancestor at the beginning of vertebrate evolution. The group L cones contain long wavelength opsins (e.g. human red, λmax 558 nm) and mid-wavelength opsins (e.g. human green, λmax 531 nm). Group S cones contain short wavelength opsins such as human blue (λmax 419 nm), chicken violet (λmax 408 nm), and ultraviolet sensitive opsins (mouse, λmax 360 nm). The group S opsins subsequently diverged to form the group M1 and group M2 cone opsin families. Group M1 cones contain blue sensitive opsins (e.g. chicken, λmax 455 nm; goldfish, λmax 441 nm), and group M2 cones contain green sensitive rod-like opsins (e.g. chicken λmax 508 nm). The group M2 family further specialised to form the group Rh family, which contains rod opsins such as human rhodopsin (λmax 500 nm) and chicken rhodopsin (λmax 503 nm). Recently an additional opsin family, group P (pinopsin, λmax 470 nm), has been identified which arose from either the group L or group M1 opsins [3,4].

In an effort to characterise the ocular photoreceptors of Atlantic salmon we amplified opsin cDNAs from salmon ocular cDNA. Our screen identified a cDNA whose conceptual translation shares only 37–42% identity to any of the known opsin families. This level of identity isolates this opsin into its own, previously unrecognised, opsin family that we tentatively call the vertebrate ancient (VA) opsins.

2. Materials and methods

2.1. Isolation

2.1.1. cDNA synthesis. Total RNA was isolated from post smolt ocular tissue from adult (>1 year) Atlantic salmon (*Salmo salaar*) reared in hatcheries at Ithaca, New York. Ocular cDNA was synthesised from 1 μg of total RNA using Superscript II reverse transcriptase and random hexamers (Life Technologies).

2.1.2. PCR. Opsin cDNA was amplified using degenerate primers BGS1 (sense) and BGS2 (antisense). BGS1 nucleotide: 5′-GTIGTITGYAARCCITTYGGIAA-3′, amino acid: VVCKPFGN; BGS2 nucleotide: 5′-TTCATRAAIACRTAIATIATIGGRTTRTA-3′, amino acid: YNPIIYVFM. Underlined nucleotides in BGS1 and BGS2 indicate a subsequent mismatch with the VA opsin sequence. The following thermal profile was performed using an MJ research PTC-100: 94° for 1 min, 55° for 1 min, 72° for 30 s, and was repeated 35 times. The remaining 5′ and 3′ coding sequences were amplified using RACE-PCR (Life Technologies) with nested VA opsin primers (BGS 3–7):

BGS3 (3′ RACE)	5′-ACCATCTAGCCTGGACCCCA-3′
BGS4 (3′ RACE nested PCR)	5′-TGCACCCGCTTTCTTCTCCA-3′
BGS5 (5′ RACE cDNA synthesis)	5′-ACCTGACGCTCCGGTT-3′
BGS6 (5′ RACE PCR)	5′-TCTTGGCGTTACCCAGCCTGCCA-3′
BGS7 (5′ RACE nested PCR)	5′-TTTCTCAGCTTCTGTAGGAGCTT-3′

Full length cDNA was amplified using primers directed against the 5′ and 3′ untranslated regions and *Pfu* DNA polymerase. All PCR products were sequenced on both DNA strands.

2.2. Phylogenetic analysis

Phylogenetic trees based on the deduced amino acid sequences of vertebrate opsins were created using the neighbour joining method [5] within the Molecular Evolution Genetic Analysis (MEGA) program. *Drosophila* RH3 opsin was used as the outgroup. The topology of the tree was confirmed by bootstrap resampling.

3. Results and discussion

We isolated a cDNA from salmon ocular tissue (Fig. 1) whose deduced amino acid translation shows only 37–42% identity to all other opsin families (Table 1). This level of identity places VA opsin into its own family. We eliminated the possibility that VA opsin belongs to seven other membrane receptor families (e.g. opsin-like *rgr* family [6], acetylcholine receptors, olfactory receptors) by demonstrating that

*Corresponding author. Fax: (44) (171) 594 5449.
E-mail: r.foster@ic.ac.uk

PII S0014-5793(97)00287-1

Full Length Coding Sequence of VA opsin

```
  1 ATG GAT ACT TTA AGA ATT GCA GTA AAT GGT GTT TCT TAT AAC GAG GCA TCA GAG ATC TAC
  1 M   D   T   L   R   I   A   V   N   G   V   S   Y   N   E   A   S   E   I   Y

 61 AAA CCA CAT GCT GAT CCG TTC ACT GGT CCA ATA ACA AAT CTG GCA CCT TGG AAT TTC GCA
 21 K   P   H   A   D   P   F   T   G   P   I   T   N   L   A   P   W   N   F   A

121 GTT TTG GCT ACC TTG ATG TTT GTT ATA ACA TCT CTG TCT CTC TTT GAA AAT TTC ACT GTT
 41 V   L   A   T   L   M   F   V   I   T   S   L   S   L   F   E   N   F   T   V

181 ATG TTG GCC ACT TAC AAG TTC AAA CAG CTG AGA CAG CCC TTA AAT TAT ATC ATA GTT AAT
 61 M   L   A   T   Y   K   F   K   Q   L   R   Q   P   L   N   Y   I   I   V   N

241 TTA TCT CTC GCT GAC TTC CTT GTG TCA CTC ACC GGT GGA ACC ATA AGT TTT CTA ACA AAC
 81 L   S   L   A   D   F   L   V   S   L   T   G   G   T   I   S   F   L   T   N

301 GCC AGA GGG TAT TTT TTC CTT GGA AAT TGG GCT TGC GTT TTG GAA GGT TTT GCA GTC ACC
101 A   R   G   Y   F   F   L   G   N   W   A   C   V   L   E   G   F   A   V   T

361 TAT TTC GGG ATT GTG GCT ATG TGG TCC CTT GCG GTC CTG TCC TTC GAG CGC TAC TTT GTG
121 Y   F   G   I   V   A   M   W   S   L   A   V   L   S   F   E   R   Y   F   V

421 ATC TGC CGG CCT CTG GGG AAT GTC CGT CTG AGG GGG AAG CAT GCA GCG CTG GGC CTG CTG
141 I   C   R   P   L   G   N   V   R   L   R   G   K   H   A   A   L   G   L   L

481 TTC GTC TGG ACC TTC TCC TTC ATC TGG ACT ATC CCT CCT GTG TTC GGC TGG TGC AGC TAC
161 F   V   W   T   F   S   F   I   W   T   I   P   P   V   F   G   W   C   S   Y

541 ACC GTC AGT AAG ATC GGC ACC ACC TGC GAG CCC AAT TGG TAT TCC AAC AAC ATT TGG AAT
181 T   V   S   K   I   G   T   T   C   E   P   N   W   Y   S   N   N   I   W   N

601 CAC ACC TAC ATC ATC ACC TTC TTT GTG ACC TGC TTC ATA ATG CCA TTG GGG ATG ATC ATC
201 H   T   Y   I   I   T   F   F   V   T   C   F   I   M   P   L   G   M   I   I

661 TAC TGC TAT GGG AAG CTC CTA CAG AAG CTG AGG AAG GTG TCC CAT GAC AGG CTG GGT AAT
221 Y   C   Y   G   K   L   L   Q   K   L   R   K   V   S   H   D   R   L   G   N

721 GCC AAG AAA CCG GAG CGT CAG GTG AGC CGC ATG GTG GTG GTG ATG ATC GTC GCT TAC CTG
241 A   K   K   P   E   R   Q   V   S   R   M   V   V   V   M   I   V   A   Y   L

781 GTG GGC TGG ACA CCC TAC GCA GCC TTC TCC ATC ATT GTC ACA GCC TGT CCC ACC ATC TAC
261 V   G   W   T   P   Y   A   A   F   S   I   I   V   T   A   C   P   T   I   Y

841 CTG GAC CCC AGA CTG GCT GCC GCA CCC GCT TTC TTC TCC AAG ACG GCT GCA GTC TAC AAC
281 L   D   P   R   L   A   A   A   P   A   F   F   S   K   T   A   A   V   Y   N

901 CCA GTA ATC TAC GTC TTC ATG AAC AAA CAG GTC TCA ACC CAA CTG AAC TGG GGA TTC TGG
301 P   V   I   Y   V   F   M   N   K   Q   V   S   T   Q   L   N   W   G   F   W

961 AGC CGC GCT TGA
321 S   R   A   *
```

Fig. 1. The full length coding sequence and conceptual translation of VA opsin are shown. Numbers indicate the position of nucleotide and amino acid residues.

they share much less identity (22%, 15%, and 13% respectively) with significant gaps and deletions in the comparison.

Salmon are tetraploid, and we have obtained a partial sequence of the second VA opsin cDNA. Of the 151 amino acid sequence overlap, we find 96% amino acid identity, with only four non-conserved amino acid substitutions. This suggests a selection pressure to conserve the amino acid sequences of both VA opsin loci. The finding of duplicate loci demonstrates that the gene was present before tetraploidy, eliminating the possibility that VA opsin is a nonfunctional, and 'drifting' correlate of a known opsin gene.

Although novel, the deduced VA opsin protein retains many conserved structural features required for photopigment function. Like all vertebrate opsins, a Kyte-Doolittle hydrophobicity plot shows that VA opsin has seven putative transmembrane domains (Fig. 2). VA opsin also contains two conserved residues intimately associated with chromophore stabilisation. First, a lysine within the seventh transmembrane

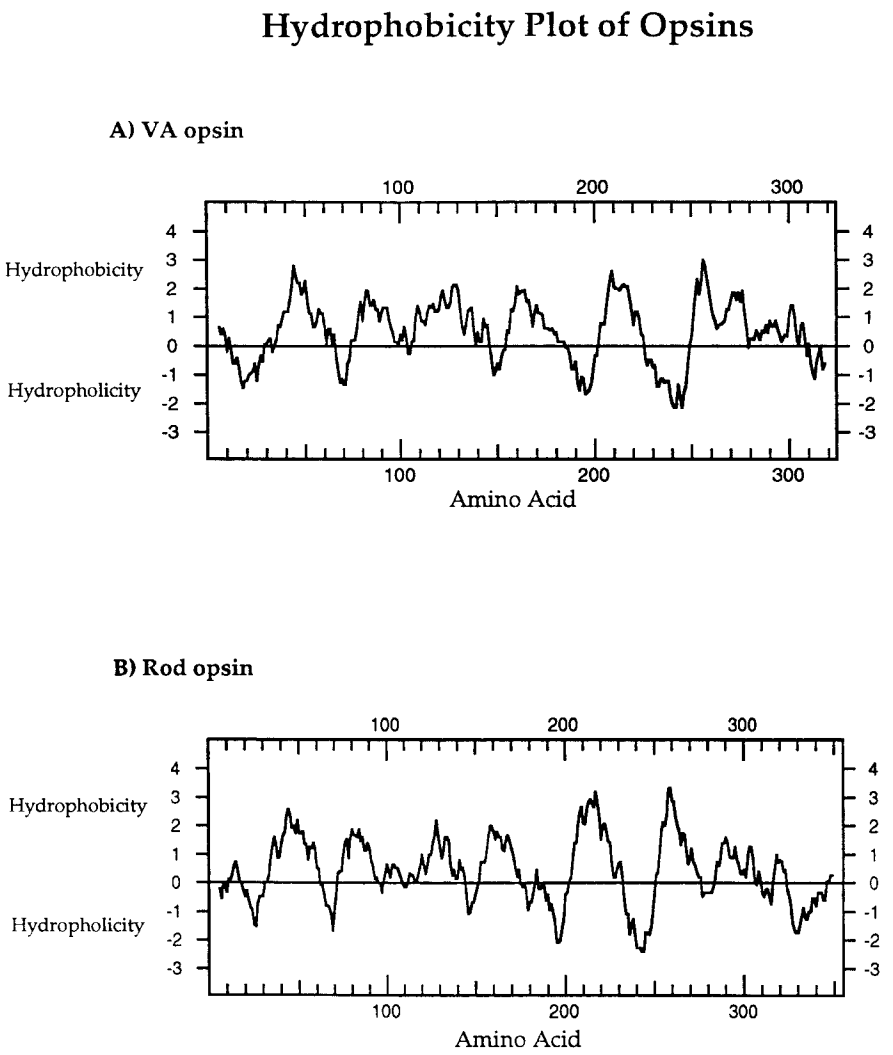

Fig. 2. Kyte-Doolittle hydrophobicity plots were generated using DNA Strider 1.2. The hydrophobicity plots shown are: (A) deduced VA opsin and (B) deduced goldfish rod opsin. Both show seven putative transmembrane domains.

domain required to bind retinal chromophore via a Schiff base linkage [7]. Second, a glutamate counterion within the third transmembrane domain which is necessary to stabilise the Schiff base linkage [8,9]. VA opsin also contains cysteines in the appropriate position to form a disulphide bond essential for correct opsin folding [10]. None of these critical residues are mutated in our partial cDNA clone of the second VA opsin. Opsins in general contain several highly conserved residues whose functions are currently not well understood. A multiple sequence analysis across the known opsin families shows 114 invariant residues. VA opsin is similarly invariant at 88% of these sites (Fig. 3).

Although VA opsin shows many conserved opsin features, several domains that are associated with phototransduction appear unique. VA opsin contains striking differences in the third cytoplasmic loop, an area which interacts with the G-protein transducin [11,12]. A comparison of VA opsin to a consensus sequence of third cytoplasmic loops reveals four non-conserved substitutions and a two amino acid deletion (Fig. 3). Additionally, it lacks conserved *N*-linked glycosylation sites which may play a role in signal transduction [13]. Finally, the termination of opsin G-protein interaction is facilitated by phosphorylation of serine and threonine residues found in the carboxyl-terminal tail [14]. The carboxyl-terminal tail of VA opsin is unique amongst the known opsins in that it is both short (approximately 14 amino acids) and has only three available residues for phosphorylation. The carboxyl tail

Table 1
Amino acid relationships of vertebrate opsin families

	P	M2	Rh	M1	L	S	VA	
P		60	61	63	63	59	59	Similarity
M2	44		75	66	60	67	53	
Rh	43	64		61	59	60	51	
M1	46	52	49		59	62	52	
L	43	44	40	40		61	55	
S	41	49	42	48	43		55	
VA	42	37	36	38	36	41		
	Identity							

Goldfish opsins (M2, Rh, M1, L, S) and chicken pinopsin (P) amino acid sequences were aligned pairwise to generate an identity and similarity comparison. Numbers below the diagonal indicate percentage amino acid identity. Numbers above the diagonal indicate the percentage similarity (identity plus conserved amino acid substitutions). VA opsin is shown in bold.

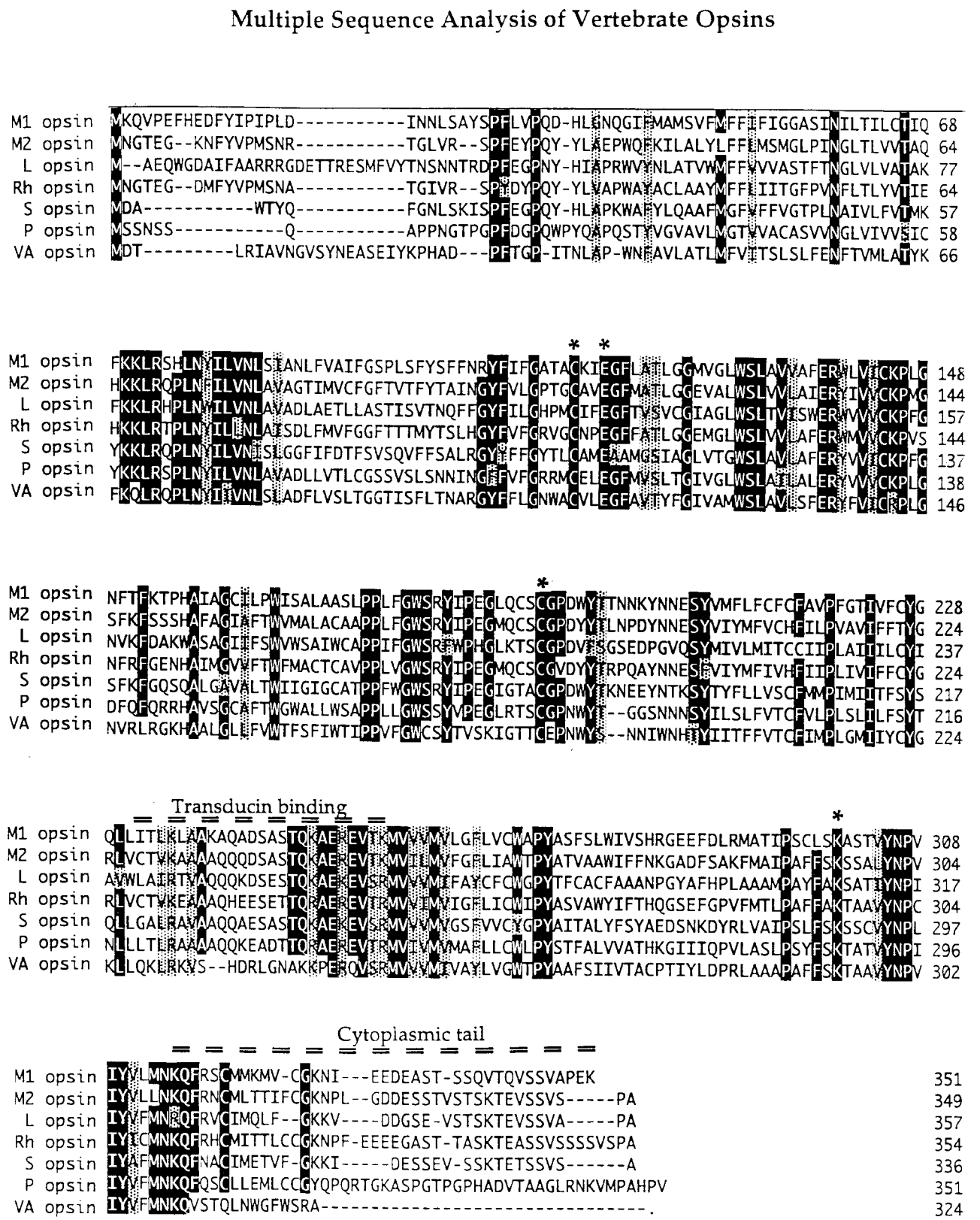

Fig. 3. Goldfish opsins (M2, Rh, M1, L, S) and chicken pinopsin (P) sequences were aligned using the Clustal method (DNAstar). Positions with at least six identical residues are highlighted in black. Positions with at least six conserved residues are highlighted in grey. The two disulphide cysteines (C), the glutamate counterion (E), and the Schiff base lysine (K) are indicated by an asterisk. A transducin binding domain and the cytoplasmic tail are beneath the bars. See text for details.

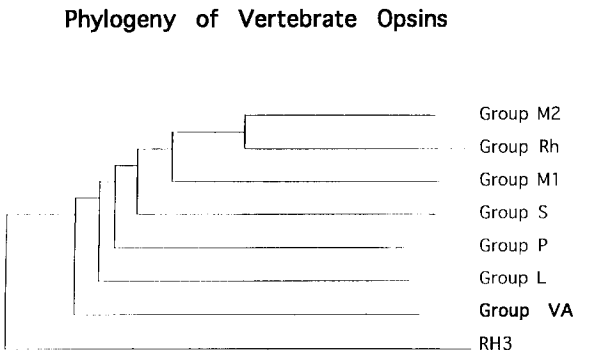

Fig. 4. A phylogenetic tree of vertebrate opsins using *Drosophila* RH3 opsin as an outgroup. The phylogenetic placement of the other opsin families agrees with previously published opsin trees [4]. Trees generated using the parsimony method (PROTPARS) yielded similar results. VA opsin appears to have diverged from a common ancestor before any of the known opsin families. Bootstrapping analysis reveals that this node is very well resolved (476 out of 500 trees).

of VA opsin also lacks palmitoylation sites (Cys-322 and Cys-323 in rod opsin). Mutants that lack these sites show reduced light-dependent phosphorylation by rhodopsin kinase [15]. A reduced number of phosphorylation sites in VA opsin may allow for extended activation of the phototransduction cascade. If so, VA opsin photopigments would be well suited to detect long-term rather than transient changes in the light environment. Whether these unique phototransduction domains in VA opsin are associated with novel phototransduction proteins remains to be determined.

Opsin isoelectric points (IEPs) were determined using the Lasergene software package (DNAstar). The deduced IEP of VA opsin is strikingly different from that of other vertebrate opsins. Cone photopigments are basic (IEP around pH 8) while rod photopigments are acidic (IEP around pH 6). The isoelectric point of VA opsin is pH 9.5, and in this respect VA opsin is unlike any known opsin family. The physiological significance of this difference is unknown but it may lend unique functional properties.

A phylogenetic analysis of VA opsin cDNA indicates that it diverged from a common ancestor before any of the known opsin families (Fig. 4). Bootstrapping analysis reveals that this node is very well resolved (476 out of 500 trees). On the basis of our phylogenetic analysis, we have assigned our novel opsin to a new family tentatively called the vertebrate ancient (VA) opsins. Preliminary Southern blot analysis suggests that VA opsins may be widely distributed among teleost fish. As a result, it seems unlikely that salmon VA opsin will remain an orphan member of the VA opsin family. We are currently expanding our search and looking for VA opsin homologues in the other classes of vertebrate.

Okano et al. [2] have argued that vertebrate cone pigments arose before rod pigments and that vertebrates acquired photopic vision before scotopic vision. If this is indeed the case, the identification and phylogenetic position of VA opsin raises the possibility that another form of light detection, pre-dating both photopic and scotopic vision, was present in the early vertebrates. We are currently exploring whether VA opsin is also expressed within the non-image forming photoreceptors (pineal, parapineal and deep brain).

The vertebrate eye does more than construct an image of the environment. Vertebrates also use their eyes to detect overall changes in the light environment, regulating functions as diverse as pupil size [16], corneal pigmentation [17], and the entrainment of endogenous biological clocks [18]. Many of these 'other' light detecting tasks remain poorly understood. We hope that the identification of the VA opsin family may resolve at least some of these uncharacterised photoreceptor functions of the eye.

Acknowledgements: This work was supported by AFOSR (91-NL-040) to R.G.F. B.G.S. is registered for a Ph.D. in the Department of Biology at the University of Virginia.

References

[1] G. Wald, Science 162 (1968) 230–239.
[2] T. Okano, D. Kojima, Y. Fukada, Y. Shichida, T. Yoshizawa, Proc Natl Acad Sci USA 89 (1992) 5932–5936.
[3] T. Okano, T. Yoshizawa, Y. Fukada, Nature 372 (1994) 94–97.
[4] M. Max, et al. Science 267 (1995) 1502–1506.
[5] N. Saitou, M. Nei, Mol Biol Evol 4 (1987) 406–425.
[6] M. Jiang, S. Pandey, H.K.W. Fong, Invest Ophthalmol Vis Sci 34 (1993) 3669–3679.
[7] J.B.C. Findlay, D.J.C. Pappin, Biochem J 238 (1986) 625–642.
[8] T.P. Sakmar, R.R. Franke, H.G. Khorana, Proc Natl Acad Sci USA 86 (1989) 8309–8913.
[9] E.A. Zhukovsky, P.R. Robinson, D.D. Oprian, Science 246 (1991) 928–930.
[10] S.S. Karnik, H.G. Khorana, J Biol Chem 265 (1990) 17520–17524.
[11] H. Kühn, P.A. Hargrave, Biochemistry 20 (1981) 2410.
[12] R.R. Franke, B. Konig, T.P. Sakmar, H.G. Khorana, K.P. Hofmann, Science 250 (1990) 123–125.
[13] S. Kaushal, K.D. Ridge, H.G. Khorana, Proc Natl Acad Sci USA 91 (1994) 4024–4028.
[14] P. Thompson, J.B. Findlay, Biochem J 220 (1984) 773–780.
[15] S.S. Karnik, K.D. Ridge, S. Bhattacharya, H.G. Khorana, Proc Natl Acad Sci USA 90 (1993) 40–44.
[16] L.J. Rubib, J.F. Nolte, Exp Eye Res 43 (1986) 771–779.
[17] S.L. Kondrashev, A.S. Khodtsev, Zool J Physiol 88 (1984) 317–325.
[18] R.G. Foster, et al. J Comp Physiol A 169 (1991) 39–50.