Functional Characterization of the Teleost Multiple Tissue (tmt) Opsin Family and Their Role in Light Detection.

Josephine K.Y. Fu

Submitted to the Department of Ophthalmology in partial fulfilment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY IN OPHTHALMOLOGY

at the

University of Oxford

Wolfson College

June 2013

Functional Characterization of the Teleost Multiple Tissue (tmt) opsin Family and Their Role in Light Detection

By Josephine Ka Yun Fu
Wolfson College, University of Oxford

Submitted to the Department of Ophthalmology on Trinity Term 2013 in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Ophthalmology

Abstract

In addition to a central circadian clock in the suprachiasmatic nucleus (SCN), zebrafish (*Danio rerio*) have local clock systems in their peripheral tissues. These peripheral tissues express a complement of clock genes that can be synchronized with the 24 h light/dark cycle and thus may be entrained by light. To date, teleost multiple tissue (*tmt*) opsin identified from *Fugu rubripes* and *Danio rerio* is the only opsin that has been proposed as a candidate to mediate this cellular photoentrainment (Moutsaki et al., 2003). Here we report the discovery of a multigene family of *tmt* opsins found not only in the teleost fishes, but in vertebrates, including amphibians, birds, reptiles, and some mammals.

Phylogenetic analysis demonstrated that this gene family consists of three main classes, tmtI, tmtII and tmtIII, with each duplicating further to give two paralogues in the zebrafish genome. Their predicted amino acid sequences contain most of the characteristic features for the function of a photopigment opsin, as well as seven transmembrane segments indicative of a G protein coupled receptor (GPCR) superfamily. Significantly, reverse transcription polymerase chain reaction (RT-PCR) reveals that the *tmt* opsin genes in zebrafish are both temporally and spatially regulated. To investigate if these tmt photopigments mediate light-activated currents in cells, each opsin was expressed *in vitro* and the responses characterised by calcium imaging, whole-cell patch clamp electrophysiology, UV-Vis spectrophotometric analysis, and bioluminescence reporter assay. Collectively, these data suggest that some of the opsin photoproteins signal via G_i-type G protein pathway. Interestingly, the spectral analysis obtained shows that most tmt opsins tested are UV-sensitive when reconstituted *in vitro* with 11-*cis* and all-*trans* retinal, indicating an intrinsic bistable dynamics. Using site directed mutagenesis on one of the tmt opsins, tmt10, the potential spectral tuning sites involved in UV detection were tested.

As part of this study, tmt opsin cDNAs were isolated from three populations of Mexican tetra (*Astyanax mexicanus*): surface, Pachon and Steinhardt. This allowed for a direct comparison between the tmt opsins present in the dark adapted species (cavefish) versus those of the light adapted species (zebrafish). It is hoped that the findings from this project will contribute to our understanding of non-visual light detection in fish and the evolution of their non-image forming photoreception.

Thesis Supervisor: Professor Mark W Hankins

Acknowledgement

First and foremost, I would like to thank my supervisor, Professor Mark W. Hankins, for giving me the support and the courage to pursue this project. Mark always had a strong belief in my abilities and his enthusiasm is what drew me to the Hankins lab in the first place. As a student, I was given incredible opportunities for which I am forever grateful.

I am also deeply grateful to all those in the Nuffield Laboratory of Ophthalmology who have been helping me throughout this project. I am indebted to Dr. Wayne Davies for teaching me all the molecular biology techniques, Dr. Steven Hughes and Dr. Laurence Brown for guiding me through the calcium imaging experiments and immunocytochemistry, Dr. Michelle McClements for western blot techniques, and Dr. Lei Zhang and Dr. Zara Melyan for the trainings in whole-cell patch clamping. Their help, patience and stimulating suggestions have assisted me over the past three years. The lab is fortunate to have Dr. Rachel Butler as facility manager for the department and Simona Dipretoro as laboratory assistant. Their work keeps the lab running smoothly and I am appreciative of their effort.

This research project could not have been possible without the great generosity of zebrafish specimens from the Whitmore lab at University College London (UCL). Especially I am obliged to Professor David Whitmore and Dr. Kathy Tamai for teaching me the dissection techniques and whole-mount in situ hybridisation. I want to thank them for all their assistance, support, interest, and valuable hints. It was also a pleasure to work alongside UCL graduates, Helen Moore and Andrew Beale. Helen taught me how to collect zebrafish embryo specimens and

provided the brain tissue sections for experimentations. Andrew has generously answered countless questions and has always been willing to give me much needed feedbacks. He also provided significant support in editing the cavefish chapter of this thesis.

Additionally, I have to thank our collaborators, Dr. Helena Bailes and Professor Robert Lucas, at the University of Manchester. They were very kind in letting me physically work in their laboratories and learn the techniques for bioluminescence reporter assay. The experience was invaluable for the development of a potential screening assay in detecting G-protein coupling activity in opsins.

Most importantly, I need to thank my family who have always believed in me. My parents, Joseph and Em-on, have taught me to be patient and determinant in pursuing my goals. My siblings, Alex and Joanne, have been incredibly supportive. They made even the hardest time of graduate school easy to endure. Finally, I am thankful to my love, Lukasz, who has always have faith in me. Thank you for reminding me what is truly important in life.

Dedicated to

My beloved parents

Fu Sai Piu Joseph and Fu Em-On

For their endless support and encouragement through my candidature.

Table of Contents

Abbreviations

aa	amino acids
aanat	arylalkylamine N-acetyltransferase
AC	adenylate cyclase
ACs	amacrine cells
ATP	adenosine trisphosphate
Am	*Astyanax mexicanus*
Bat-1	bovine AT-rich sequence-1
BCs	bipolar cells
BMAL1	brain and muscle aryl hydrocarbon receptor nuclear translocator (ARNT)-like 1
BRE	B Recognition Element
Bt	*Bos taurus*
$[Ca^{2+}]_i$	intracellular calcium concentration
CC	crista cerebellaris
CCe	corpus cerebella
cDNA	complementary DNA
ceh-10	*Caenorhabditis elegans* homeobox-10
Ci opsin	ciona opsin
Cinf	comissura infima of Haller
CK1	casein kinase 1
CL	cytoplasmic loop
CLOCK	circadian locomotor output cycles kaput
CNS	central nervous system
CNG channels	cyclic-nucleotide gated channels
CRE	cAMP response element
CREB	cAMP-response element binding protein
CRY	cryptochrome
crx	cone-rod homeobox
CTCF	CCCTC-binding factor
Cys	cysteine
DAG	diacylglycerol
DDM	n-Dodecyl-β-D-Maltoside ULTROL® Grade
DIG	digoxgenin

D-box	destruction-box
DPE	Downstream Promoter Element
dpf	days post fertilisation
Dr	*Danio rerio*
E-box	enhancer-box
EC	extracellular loop
ECL	external cellular layer
EG	eminentia granularis
Elk-1	E26-like transcription factor 1
EP3	prostaglandin E_2 receptor subtype
EtBr	ethidium bromide
ER	endoplasmic reticulum
EST database	expressed sequence tag database
Exo-rod opsin	extraretinal rod-like opsin
F	forward primers
FBS	foetal bovine serum
G	gut
G protein	guanine nucleotide-binding protein
GPCR	G protein coupled receptor
G_t	transducin
Hd	periventricular hypothalamus
HEK 293T	human embryonic kidney 293
HINGS	heat-inactivated goat serum
HM	hybridisation mix
HMM	Hidden Markov Model
HNF	hepatocyte nuclear factors
hOPN4	human melanopsin
hpf	hours post fertilisation
HPLC	high performance liquid chromatography
ICC	Immunocytochemistry
IRBP	interphotoreceptor retinoid-binding G-protein
INR	initiator element
IPL	inner plexiform layer
IPTG	isopropyl β-D-1-thiogalactopyranoside
IS layer	inner segment layer

JTT	Jones-Taylor-Thornton substitution matrix
L	liver
LCa	lobus caudalis cerebella
LDH	lactate dehydrogenase
LOT	lateral olfactory tract
LRAT	lecithin-retinol acyltransferase
LWS opsin	long-wave sensitive opsin
LVII	lobus facialis
LX	lobus vagus
M	medulla oblongata
MAB	maleic acid buffer
mAChR1	muscarinic acetylcholine receptor
MCL	Maximum Composite Likelihood
MeOH	methanol
miRNA	micro RNA
MWS opsin	middle-wavelength-sensitive opsin
MYA	million years ago
NC	notochord
Neuro-2A cells	mouse neuroblastoma cells
NIF	non-image forming photoreception
nr2e3	nuclear receptor subfamily 2 group E member 3
nrl	neural retina leucine zipper
Opn1	cone opsin
Opn2	rod opsin
Opn3	panopsin/encephalopsin
Opn4	melanopsin
Opn5	neuropsin
ORF	open reading frame
Otx	orthodenticle homeobox
P-opsin	pinopsin
PP opsin	parapinopsin
PBS	phosphate-buffered saline
PC	Pachón
Pce-1	photoreceptor conserved element-1
Per	Period

PFA	paraformaldehyde
PGm	medial periglomerular nuclei
PGZ	periventricular gray zone
PLC	phospholipase C
PTN	posterior tuberal nuclei
PTU	1-phenyl 2-thiourea
PVDF	polyvinylidene
R1	reverse primers for putative long isoforms
R2	reverse primers for putative short isoforms
RA	retinoic acid
Ret-1	retina-specific region-1
RGCs	retinal ganglion cells
RGR opsin	RPE-retinal G-protein-coupled receptor opsin
RH1	rod opsin
Rho 1D4	bovine rhodopsin ID4 epitope tag
Rhod-2 AM	rhodamine-2 acetoxymethyl
R-ISH	RNA *in situ* hybridisation
RLU	raw luminescence units
RNAi	RNA interference
ROIs	regions of interest
ror$_\beta$	rar-related orphan receptor beta
RP4	retinitis pigmentosa type 4
RPE	retinal pigment epithelium
RPE 65	retinal pigment epithelial protein 65
RRH	RPE-derived rhodopsin homologue
RT-PCR	reverse transcription polymerase chain reaction
SCN	suprachiasmatic nuclei
S.E.M.	standard error of the mean
SF	Surface
SH	Steinhardt
SSC	saline-sodium citrate
SWS1 opsin	short-wave sensitive type 1 opsin
SWS2 opsin	short-wave sensitive type 2 opsin
T	neurocranial trabecula
TBST	Tris-buffered saline with Tween-20

TEF	thyrotrope embryonic factor
TFBS	transcription factor binding sites
TGN	tertiary gustatory nucleus
tmt opsin	teleost-multiple tissue opsin
TM	transmembrane segments
TSS	transcription start site
TRP channel	transient receptor potential channel
U	units
UTR	untranslated region
UV	ultraviolet
UVS	ultraviolet-sensitive
VA opsin	vertebrate ancient opsin
VAL	long isoform of vertebrate ancient opsin
VAM	medium isoform of vertebrate ancient opsin
VAS	short isoform of vertebrate ancient opsin
λ_{max}	absorption maxima
ZT3	zeitgeber 3

CHAPTER 1

Introduction:

Non-visual opsin-mediated Irradiance Detection

CHAPTER 1: Introduction

1.1 Non-image-forming photoreception and photoreceptors

Majority of the life forms on Earth rely on sunlight for survival, photoreception is therefore vital for most organisms. In animals, the process of light detection is mediated by photoreceptor cells which contain light-absorbing pigment molecules. These pigments are composed of an opsin protein coupled to a covalent ligand chromophore, typically a 11-*cis* or an all-*trans* isoform of vitamin A retinaldehyde (Wald, 1968). Opsins are members of the G-protein coupled receptor (GPCR) superfamily, characterised by a conserved lysine residue in the seventh transmembrane domain for attachment of retinal chromophore (Menon et al., 2001). They function to activate light-induced signalling cascades. Although light absorption by these photopigments plays a key role in cellular photoreception, this is only the initial step in the whole cascade of phototransduction. Transmission of light signal into the cell can mediate intracellular signalling pathways involving other downstream effector proteins and drive various biological activities (Choe et al., 2011). Such photoreception system can occur at a single-cell level, as exemplified by light-sensitive *Dictyostelium discoideum* amoeba (Schlenkrich et al., 1995), to a collection of irradiance detection patches forming pigmented eye spots (ocelli) in planarians (Brown et al., 1968, Santillo et al., 2006), to the complex eyes observed in higher animals.

In the eyes of vertebrates, they have a duplex retina of cone and rod photoreceptors that mediate image-formation with luminescence and chromatic

information (Land and Fernald, 1992). The visual opsins that are present in these photoreceptors can be classified into five different subtypes based on their phylogenetic relationships and absorbance maxima: (1) long-wavelength-sensitive (LWS) cone opsin, (2) short-wavelength-sensitive type 1 (SWS1) cone opsin, (3) short-wavelength-sensitive type 2 (SWS2) cone opsin, (4) middle-wavelength-sensitive (MWS or RH2) cone opsin, and (5) rod-opsin (opn2 or RH1) (Yokoyama, 2000, Davies, 2011). Over the past two decades, other vertebrate non-visual opsins have also been discovered from retinal and extra-retinal sites, many of which are still functionally unknown. These include: retinal G-protein coupled receptor (RGR), peropsin, pinopsin, parapinopsin, parietopsin, exo-rod opsin, vertebrate ancient (VA) opsin, panopsin/encephalopsin (opn3), teleost multiple tissue (tmt) opsin, melanopsin (opn4) and neuropsin (opn5) (reviewed by Peirson et al., 2009, Terakita, 2005, Kojima and Fukada, 1999). Melanopsin and VA opsins are examples of non-visual opsins that are well characterised (reviewed by Foster and Hankins, 2002, Davies et al., 2010), and they have been shown to control a variety of non-image-forming (NIF) physiological functions including: regulation of circadian photoentrainment, synchronising an organism's 24-h biological activity to environmental light and dark cycle (Brown and Robinson, 2004, Davies et al., 2010); skin pigmentation, which is colouration of tissues by deposition of pigments in the cells (Provencio et al., 1998); and pupillary light reflex, which is regulation of pupil constriction in response to light that enters the eye (Gooley et al., 2012). It was traditionally thought that NIF photoreception is specifically carried out by non-cone, non-rod photoreceptors expressing these non–visual opsins (Foster et al., 2006).

However, recent evidence shows that both cones and rods contribute to circadian photoentrainment by sending signals to melanopsin-expressing retinal ganglion cells (RGCs) at high and low light intensity respectively (Altimus et al., 2010, Lall et al., 2010). It is important to note that there also exists a number of visual and non-visual opsins in the invertebrates, such as ciona (Ci) opsins in the ascidian chordate (*Ciona intestinalis*) (Kusakabe et al., 2001), arthropod colour visual rhodopsins (Koyanagi et al., 2008), jellyfish opsins (Suga et al., 2008), coral opsins (Mason et al., 2012), echinoderm opsins (Lesser et al., 2011), those from the lancelet (also known as amphioxus) (Holland et al., 2008, Vopalensky et al., 2012) as well as the honey bee opsins (Velarde et al., 2005, Townson et al., 1998). The presence of such a diverse repertoire of opsins in both invertebrates and vertebrates demonstrates that besides vision, photoreception is linked to many non-visual aspects of animal physiology.

Although some non-visual opsins have been studied extensively, our understandings of how NIF photoreception functions, in particularly outside of the retina, are still unclear. Of all the opsins discovered thus far, the tmt opsins identified from zebrafish (*Danio rerio*) has the most widespread pattern of expression in CNS as well as in peripheral tissues (Moutsaki et al., 2003). Its profile of expression closely matches that of the peripheral circadian oscillators, which are entrainable by light. These data, coupled with phylogenetic evidence and its key structural features indicative of a photopigment, strongly suggest that tmt opsin may be responsible for mediating NIF photoreception functions such as circadian photoentrainment. The main aim of this project was therefore to characterise the *tmt opsin* gene family and to express it functionally for investigation of its potential role(s) in NIF photoreception.

1.1.1 Light sensing and the phototransduction cascade

Light detection by photopigments involves the absorption of photons, which converts the GPCR molecules into an active state. Upon binding specific guanine nucleotide-binding G-protein (G-protein), the activated photopigments trigger a signalling cascade in the cell that results in generation of electrical signals. This process is known as the phototransduction cascade. In both vertebrates and invertebrates, non-visual photoreceptors generally share the same components with ocular photoreceptors, involving a photopigment (opsin protein + retinal chromophore molecule), and similar G-protein signalling pathways that are coupled to different secondary messengers for mediating specific functional outputs (Wolken and Mogus, 1979). The first two steps of phototransduction are illustrated in Fig. 1.1: (1) photon absorption, and (2) G-protein activation.

Photopigments absorb photons through the bound retinal, which is a vitamin A-based aldehyde covalently linked by Schiff base to a conserved lysine residue (Lys296) in the opsin (Palczewski et al., 2000). There exist different isomers of retinal chromophore (e.g. all-*trans*, 9-*cis*, 11-*cis*, and 13-*cis*) and its presence is essential for light sensitivity of the pigment molecule. The *cis*-retinal has an angulated structure with hydrogen (H) atoms located on one side of the double bond throughout the entire chromophore, whereas the *trans*-retinal is a straight structure with hydrogens on opposite sides of the double bond (see Fig. 1.2).

Vertebrate visual photopigments are commonly coupled to 11-*cis* retinal as their endogenous chromophore. Absorption of light by the chromophore causes 11-*cis* retinal to photoisomerise into an all-*trans* configuration (Futterman and

Futterman, 1974; Fig. 1.2A). This in turn leads to a series of changes in the protein moiety of the photopigment, involving movement of transmembrane segments (TM5 and 6) that exposes a G-protein binding site (Ridge and Palczewski, 2007, Scheerer et al., 2008). Attachment of transducin (G_t) to the visual opsins would trigger a signalling cascade that leads to the closing of cyclic-nucleotide gated (CNG) channels, eventually generating a hyperpolarising response in the photoreceptor cell (Arnis and Hofmann, 1993). Simultaneously, conformational changes in the pigment increase accessibility for water molecules to enter and hydrolyse the Schiff based linkage between the chromophore and the opsin as all-*trans* retinal is generated (Ridge and Palczewski, 2007). Since all-*trans* retinal is not able to form a stable association with the opsin, it is released from an aperture between TM1 and TM7 of the retinal binding site as all-*trans* retinol, 'bleaching' the photopigment (Scheerer et al., 2008). The free all-*trans*-retinol is then transported by interphotoreceptor retinoid-binding G-protein (IRBP) to the retinal pigment epithelium (RPE) to undergo a series of enzymatic regeneration steps to be converted back to 11-*cis*-retinal in a process called the visual cycle (see Fig. 1.1) (Ala-Laurila et al., 2006, Jin et al., 2009, Qtaishat et al., 2005). The bleached opsin binds to an arrestin protein that causes a conformational change to its inactive dark state (Vishnivetskiy et al., 2007). Finally, evidence has suggested that 11-*cis* retinal might be transported by IRBP to the photoreceptor outer segment membrane (Gonzalez-Fernandez, 2012) and conjugates with an inactive opsin to regenerate a new photopigment, maintaining visual sensitivity. Since restoration of photosensitivity in these bleached visual pigments requires a new supply of 11-*cis* retinal, this defines their monostable characteristics.

In the non-visual system of vertebrates, however, much less is known about their chromophore identity and the process of chromophore regeneration. It has been suggested that certain non-visual opsins, such as opn4 and opn5, may form bistable photopigments that bind both 11-*cis* and all-*trans* retinal in dark and light-activated states (Mure et al., 2009, Melyan et al., 2005, Lucas, 2006, Yamashita et al., 2010). These bistable pigments possess inherent photo-reversal capability for the regeneration of all-*trans* retinal to 11-*cis* retinal using specific wavelength of light (Fig. 1.2B,Hillman et al., 1983, Mure et al., 2007). Unlike monostable pigments, which have been postulated to release the all-*trans* retinal upon activation, the structure of bistable pigments probably do not allow for an exit point. Hence, the pigment can be continuously activated by light without being bleached. Such bistable property is also observed in invertebrate opsins such as Arthropod and Cephalopod visual pigments (Tsukamoto and Terakita, 2010). Vertebrate non-visual pigments couple to a more diverse range of G-proteins compared to the visual pigments. For example, opn5 has been shown to bind G_i protein (Kojima et al., 2011), whereas opn4 activates G_q, G_{11} or G_{14} protein (Sexton et al., 2012). Interestingly, the G_q signalling pathway activated by opn4 and the cellular response generated broadly resemble those in invertebrate photoreceptors. Phototransduction of opn4 and invertebrate opsins both results in the opening of transient receptor potential (TRP) channels and triggers Ca^{2+}/Na^{+} influx, leading to a depolarizing response in photoreceptor cells (Hartwick et al., 2007, Hankins et al., 2008). This suggests that perhaps opn4 is evolutionally and functionally more closely related to the invertebrate opsins than to vertebrate opsins.

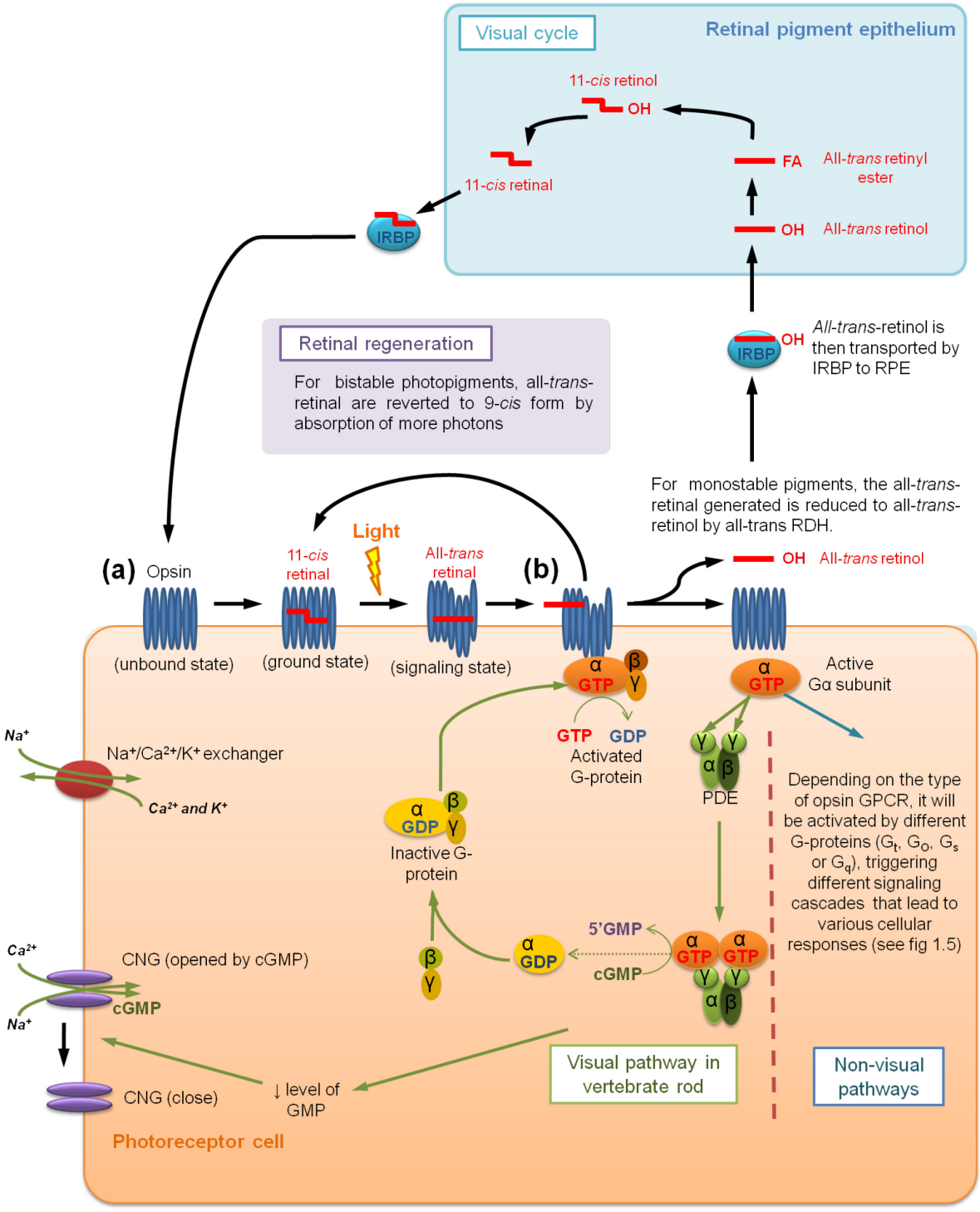

Visual cycle
Retinal pigment epithelium
11-cis retinol
OH
11-cis retinal
FA
All-trans retinyl ester
OH
All-trans retinol
IRBP
IRBP
OH
All-trans-retinol is then transported by IRBP to RPE
Retinal regeneration
For bistable photopigments, all-trans-retinal are reverted to 9-cis form by absorption of more photons
For monostable pigments, the all-trans-retinal generated is reduced to all-trans-retinol by all-trans RDH.
OH
All-trans retinol
Light
(a)
Opsin
11-cis retinal
All-trans retinal
(b)
(unbound state)
(ground state)
(signaling state)
α GTP
β γ
Active Gα subunit
GTP GDP
Activated G-protein
Na+
Na+/Ca2+/K+ exchanger
Ca2+ and K+
α GDP
β γ
Inactive G-protein
γ γ
α β
PDE
Depending on the type of opsin GPCR, it will be activated by different G-proteins (Gt, Go, Gs or Gq), triggering different signaling cascades that lead to various cellular responses (see fig 1.5)
5'GMP
cGMP
α GDP
β γ
α GTP
α GTP
γ γ
α β
Ca2+
CNG (opened by cGMP)
cGMP
Na+
CNG (close)
↓ level of GMP
Visual pathway in vertebrate rod
Non-visual pathways
Photoreceptor cell

Figure 1.1 General mechanism of light-induced signalling in vertebrate visual opsins as an example of phototransduction: **(a)** Light is captured by 11-*cis* retinal that is bound to a visual opsin in photoreceptor cells. Absorption of light isomerises the retinal into all-*trans* configuration and simultaneously exposes a G-protein binding site in the opsin molecule. **(b)** In light-activated state, specific heterotrimeric G-protein (composed of $G_{\alpha,}$ G_{β} and G_{γ} subunits) can bind to the opsin, which catalyses an exchange reaction of GDP to GTP. The GTP-bound G_{α} subunit then dissociates from the $G_{\beta\gamma}$ subunits, exposing an active site that can act on specific downstream effectors. In vertebrate cones and rods, active G_{α} subunit activates photodiesterase (PDE6) which breaks the phosphodiester bond of cGMP and give 5'GMP. PDE is a superfamily of enzymes that comprises at least eleven different isoforms (Halpin, 2008). Different isomers of PDE have different kinetics, substrate specificity and tissue specificity, regulating a broad range of physiological functions. The reaction catalysed by PDE6 decreases the intracellular level of cGMP and causes the closure of cyclic nucleotide gated (CNG) channels, leading to hyperpolarization of the photoreceptor cell. The free all-*trans* retinal that is released from the activated pigment is transported by interphotoreceptor retinoid-binding G-protein (IRBP) to retinal pigment epithelium (RPE), where it goes through the visual cycle for conversion into 11-*cis* retinal. In the RPE, the all-*trans* retinol is esterified by lecithin-retinol acyltransferase (LRAT) into all-*trans* retinyl ester and then hydrolysed by isomerohydrolase RPE 65 to 11-*cis* retinol. 11-*cis* dehydrogenase (11-*cis* RDH) oxidises the 11-*cis* retinol into 11-*cis* retinal before it is being transported by IRBP back into visual photoreceptor cells and conjugate with an apo-opsin to form active visual pigment. The illustration was drawn based on explanations by Shichida and Matsuyama (2009), and von Lintig et al. (2010).

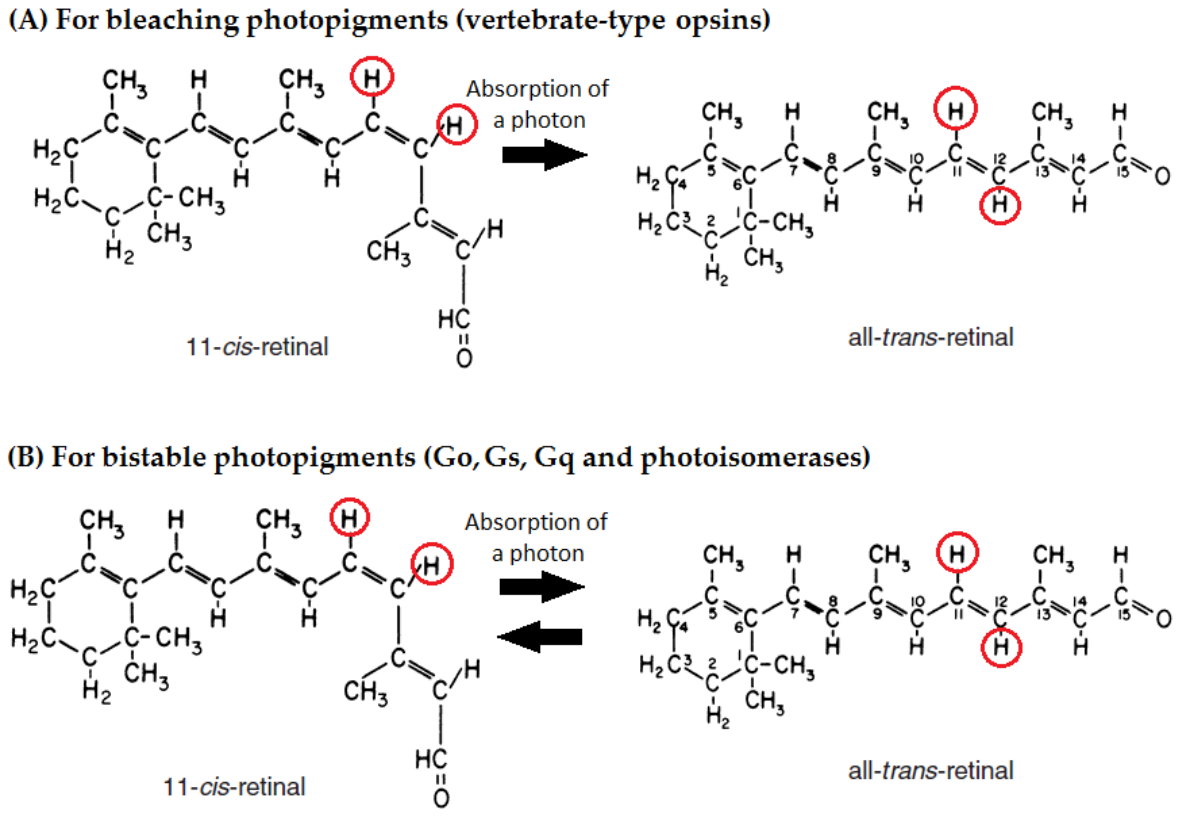


Figure 1.2 Diagram showing the conversion between 11-*cis*-retinal and all-*trans*-retinal for **(A)** monostable photopigments and **(B)** bistable photopigments. In 11-*cis*-retinal, hydrogen molecules (circled red) are located on the same side of the double bond, whereas all-*trans*-retinal has hydrogen molecules located on opposite sides of the double bond (Terakita, 2005).

1.2 The retinal chromophore

Apart from retinal, there are actually in total four different types of chromophores that are used for formation of all photopigments: (1) retinal (also known as A1), (2) 3,4-dehydroretinal (A2), (3) 3-hydroxyretinal (A3), and (4) 4-hydroxyretinal (A4) (Shichida and Matsuyama, 2009). These isoforms differ in the location of H molecules within the chromophore structure, the number of double bonds in the ring, as well as the presence of a hydroxide group (Fig. 1.3).

Although vitamin A-based retinal functions as universal light sensing chromophore, different animal species have adapted to use specific isomers of the retinal. Opsins that are associated with A1 chromophore are called rhodopsins, and this type of pigments is commonly found across all vertebrates and invertebrates (Sakmar et al., 2002). The absorption maxima (λ_{max}) of rhodopsin lie over a range between 345 nm (UV) to 575 nm (visible) depending on species specificity (Sen et al., 1984). Opsins which are coupled to A2 chromophores are known as porphyropsins, and they are predominantly found in aquatic animals including freshwater fish, amphibians, and reptiles (Zhong et al., 2012). The λ_{max} of porphyropsins are distinctly more red-shifted compared to rhodopsins, and magnitudes of the shift are dependent on the initial λ_{max} of the A1 retinal system (Warrant and Nilsson, 2006). That is, change in bandwidth is the smallest in pigments which have λ_{max} in the UV-range compared to those in the visible region. This red-shifted property in porphyropsin is due the structure of A2 retinal, which has an extended π-system with an additional double bond inside the ring, thus allowing it to absorb light at longer wavelengths compared to A1 retinal (Dartnall and Lythgoe, 1965, Harosi, 1994).

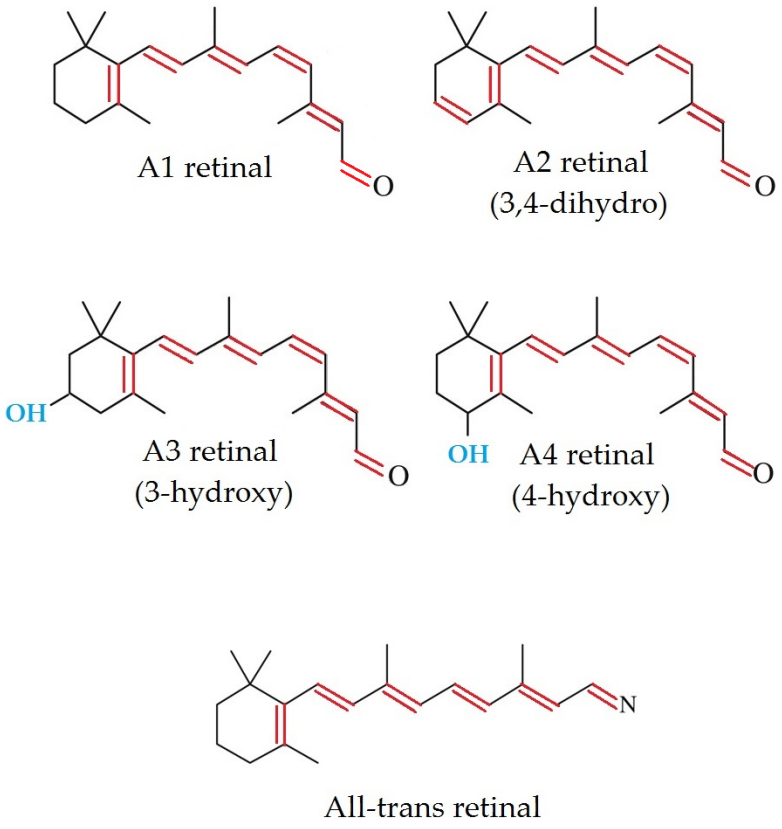


Figure 1.3 Schematic diagrams showing the molecular structure of A1, A2, A3, and A4 11-*cis* retinal, as well as all-*trans* retinal (Shichida and Matsuyama, 2009). Double bonds are highlighted in red, and hydroxide groups are labelled in blue.

The differences in spectral absorbance between rhodopsin and porphyropsin are closely related to the ecology habitat of the species, and comparison is often made between freshwater and deep-sea marine fish. In deep ocean, light at around 480 nm (blue) penetrates furthest through the water (Jerlov, 1976), whereas in fresh water habitats, light at longer wavelengths 550 nm (yellow) and 650 nm (red) dominate (Reckel et al., 2002, Leech and Johnsen, 2009). It has also been shown that there are greater variations in luminescence level between day and night in the marine ocean in comparison with shallow, freshwater habitats (Collin and Marshall, 2003). In order to have spectral sensitivity specifically tuned to match the spectrum of light in these different water environments, deep-sea fish have adapted to use rhodopsins whilst freshwater fish have the red-shifted porphyropsins as their primary pigments to enhance visual sensitivity (Wald, 1939, Toyama et al., 2008). Some migratory species, such as lamprey *(Petromyzon marinus)*, European eel *(Anguilla anguilla)*, and Atlantic salmon *(Salmo salar*), are known to shift between A1 and A2 chromophore in their visual system through development or under different light environment during seasonal changes (Wald, 1958, Collin et al., 2003a, Wood et al., 1992, McFarland and Allen, 1977). Research has shown that such optimisation of the visual system in animals closely matching their environments has biological significance in visual communication, mating choices, navigation, and searching for food resources (Leech and Johnsen, 2009, Crescitelli et al., 1985). If a similar chromophore exchange system occurs for the non-visual opsins, spectral tuning in this way may be advantageous.

The other two *cis* isomers of retinal chromophore, A3 and A4, are found in invertebrate species. Most insects use A1 and/or A3 retinals, which is believed to

contribute to their colour and ultraviolet vision (Vogt, 1989, Seki and Vogt, 1998). A4 retinal has so far only been observed in firefly squid (*Watasenia scintillans*), together with A1 and A2 retinals. Using three different forms of retinals, firefly squid possesses a range of photopigments with different absorbance maxima that constitute their colour vision (Seidou et al., 1990, Shichida and Matsuyama, 2009). Given the presence of different chromophores and retinal subclasses, invertebrates can alter their spectral sensitivity like some vertebrates (as described previously in aquatic animals) for detecting light in various environments. The results of diverse variations in photopigment complements are likely to provide the animals significant adaptive advantages in gaining the highest possible visual sensitivity.

1.3 The opsin superfamily

Opsins are a large family of 30-60 kDa transmembrane GPCR proteins that function as light sensors in photoreceptor cells of most species (Miyaguchi et al., 1992). Depending on specific chemical interaction between the chromophore and the opsin protein, the photopigment molecule can absorb light at certain wavelengths. Since peak sensitivities and their spectral shifts are limited by only four chromophore types, the opsin is the main way of changing the absorbance profile of the photopigment. This is achieved by variations in the structure of different opsins, which are determined by their genetic and amino acid sequences (Carroll and Jacobs, 2008, Yokoyama, 2008). With many possible versions of the opsin proteins available, there are correspondingly many different types of photopigments encompassing a broad range of absorption spectrum.

1.3.1 Classification

The opsin superfamily was traditionally divided into two major classes based on their expression in different types of photoreceptors: rhabdomeric (r-type) and ciliary (c-type) opsins (Arendt, 2003, see Fig. 1.4). Although both cell types employ G-proteins to mediate phototransduction, the specific G-proteins required and the biochemical cascades involved can vary. In general, ciliary opsins signal via activation of G_t, G_i, G_o families of G-proteins and alter intracellular level of cyclic nucleotides (Terakita, 2005), whereas invertebrate opsins and opn4 use the G_q family of the G-proteins, involving phospholipase C (PLC), phosphoinositol and TRP channels (Hardie and Raghu, 2001). However, with the discovery of numerous new opsins, this simple classification of just r- and c-type opsin groups does not encompass the whole opsin family. There is an additional smaller class of opsins, the photoisomerase-like opsins (e.g. retinochromes and RGR opsins), which do not seem to couple with any G-protein and therefore are unlikely to have a function in signal transduction (Shichida and Matsuyama, 2009). Instead, these photoisomerases have the ability to bind all-*trans* retinal and convert them into 11-*cis* form upon irradiation. Studies have shown that with RGR opsins, they may also act as a cofactor that can interact directly or indirectly via other components with retinal pigment epithelial protein 65 (RPE65), an essential protein that helps to convert all-*trans* retinal into 11-*cis* retinal (Chen et al., 2001, Bhattacharya et al., 2002, Chen et al., 2005). The involvement of RGR in synthesis, mobilization, as well as degradation of all-*trans* retinyl ester in RPE facilitates the processing of retinoid in the visual cycle (Wenzel et al., 2005, Radu et al., 2008).

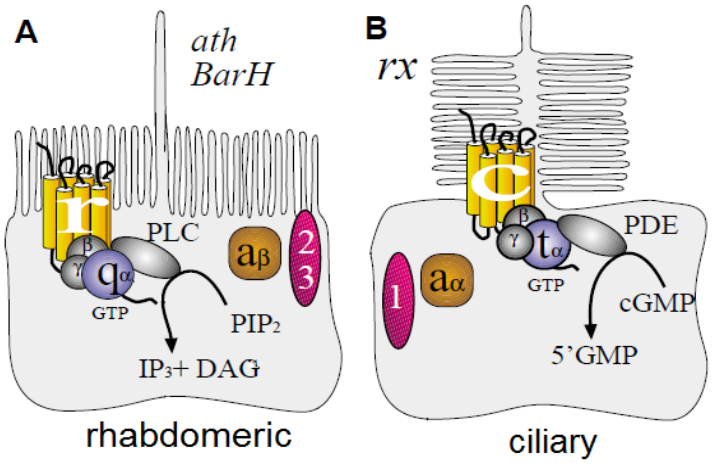


Figure 1.4 (A) Rhabdomeric and **(B)** Ciliary photoreceptors, with distinctive morphology and G-protein coupled signaling cascades (Arendt, 2003).

1.3.2 Genome phylogeny and evolution

To date, thousands of opsin genes have been sequenced from a wide range of organisms from bacteria to vertebrates, highlighting the immense diversity in the opsin family. Molecular phylogenetic analysis shows that this gene family in animals can be categorized into at least seven subfamilies: the vertebrate G_t-coupled visual/non-visual opsin subfamily; the encephalopsin/tmt-opsin subfamily; the G_o-coupled opsin subfamily; the G_s-coupled opsin subfamily (restricted to jellyfishes); the G_q-coupled opsin subfamily; the retinal photoisomerase subfamily; and the neuropsin subfamily (see Fig. 1.5). There is less than 20% identity at the amino acid sequence level shared between these subfamilies, but greater than 40% identity between members within each family (Shichida and Matsuyama, 2009). Each of the opsin subfamilies has evolved to use different phototransduction signalling pathways, as summarized in Fig. 1.5, and this functional divergence coincides with division of the subfamily.

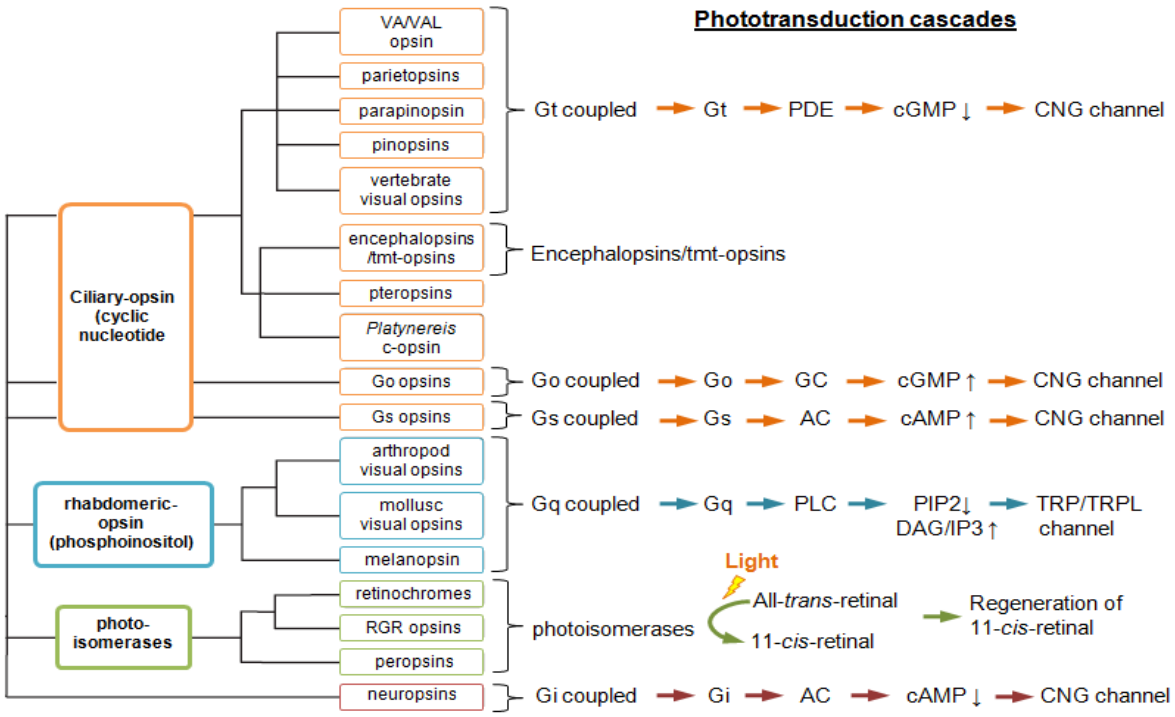


Figure 1.5 A schematic representation of the phylogenetic relationships and functional diversity of the opsin superfamily. Opsins are broadly divided into ciliary and rhabdomeric opsins. In addition, there are two smaller classes, the photoisomerases and the neuropsins, which do not signal via the G-protein signalling cascade. These opsins can be further subdivided into seven groups, with distinctive phototransduction pathways. In general, the ciliary opsins signal via a cyclic nucleotide signaling pathway, whilst the rhabdomeric opsins act on the phosphoinositol signaling pathway. As for the photoisomerases, they can form bistable pigments with chromophore and catalyse the regeneration of 11-*cis*-retinal. The phototransduction mechanism of neuropsin has recently shown to be a Gi-protein coupled pathway (modified from Shichida and Matsuyama, 2009, Yamashita et al., 2010).

1.3.3 Characteristic structural features and functions of opsin protein

The opsin is composed of a single polypeptide chain of ~350-400 amino acids. Like all other GPCRs, it is folded into seven helical segments (TM1-7) threading across the membrane (see Fig. 1.6). These transmembrane segments are linked together by three extracellular (EC1-3) and three cytoplasmic loops (CL1-3), with the amino (N-) terminal domain located in the extracellular side of the photoreceptor cell and the carboxyl (C-) terminal domain on the cytoplasmic side (Sakmar, 2002). At the beginning of C-terminus, there is a kink structure located just after TM7 forming helix 8 (H8), which is positioned on the intracellular surface of the lipid bilayer.

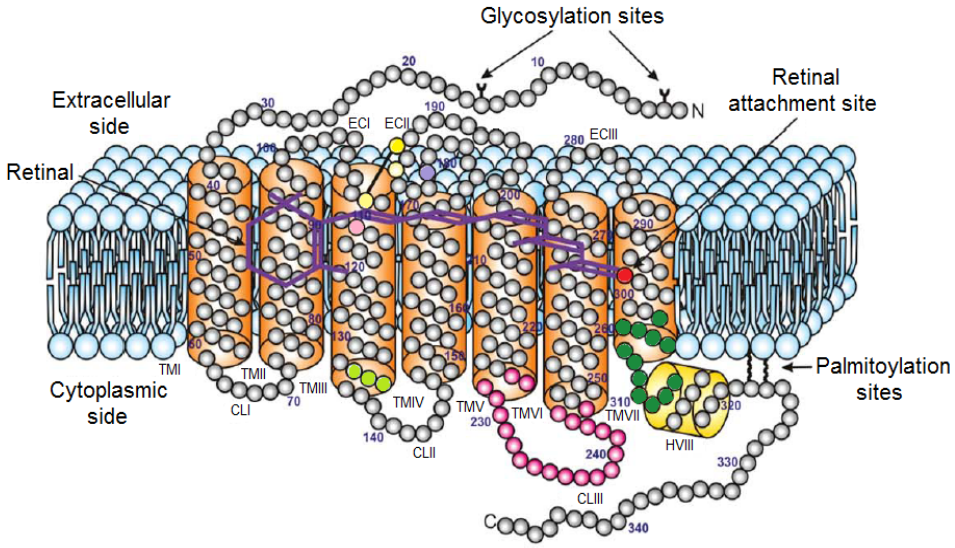


Figure 1.6 Schematic diagram of a representative photopigment drawn based on the crystal structure of bovine rod opsin (Davies et al., 2010). The opsin protein has seven helical transmembrane domains (TMI-VII) (orange), with 3 cytoplasmic loops (CLs) and 3 extracellular loops (ECs). The 11-*cis*-retinal is shown in dark purple. Amino acids that are conserved throughout the opsin family are in coloured circles: the conserved Cys110 (pale yellow) and 187 (dark yellow), a retinal-attachment site (Lys296) (red), a counterion (Glu113 or Tyr113) (pale pink), the displacement of counterion to Glu181 in many non-visual opsins (pale purple), the NPxxY(xx)5,6F motif (302-313) (dark green), and a conserved Glu134 within an ERY motif (134-136) (pale green) .

The generation of a high resolution crystal structure and multiple mutational studies of a well characterised opsin, the bovine rod opsin (RH1), have revealed many of the critical residues that are required for correct folding and function. Two highly conserved cysteine residues (Cys110 and Cys187) have been shown to provide a disulphide bond that restricts the movement of the β4 strand within EC2 (Karnik et al., 1988). It is postulated by Karnik et al. (1988) that this β4 strand and EC2 loop act as a cap to trap the retinaldehyde chromophore in its binding pocket within the opsin. The region was later confirmed by mutational studies to be important for stabilizing the pigment as well as for photobleaching function (Doi et al., 1990, Yan et al., 2002, Davidson et al., 1994). The amino acid sequence of β4 strand is highly conserved amongst vertebrate opsins (Menon et al., 2001), whilst that of the EC2 loop is more variable (Bradshaw and Dennis, 2004). It is possible that variability in the sequence of EC2 loop prevents the release of retinal after photoisomerisation in bistable pigments. The chromophore itself is covalently attached to a conserved, positively charged lysine residue (Lys296) in TM7 via a Schiff base linkage, which is stabilised by a negatively charged counterion (Pepe, 1999). In vertebrate visual opsins and some non-visual opsins (e.g. VA opsin and opn3), a highly conserved glutamate (Glu113), or sometimes aspartate (D113) in TM3 serves as the negative counterion to this site (Nathans, 1990). In many non-visual opsins, however, position 113 is occupied by an uncharged or partly positively-charged amino acid (e.g. tyrosine in opn4, histidine in RGR, and glutamine in parietopsin), with the counterion probably displaced to a highly conserved Glu181 (also negatively charged) instead. There are exceptions to this counterion switch model, with all LWS cone pigments possessing a histidine at homologous site 181 that is involved in

spectral tuning (Birge and Knox, 2003), and VA opsin carrying a serine residue at the same site but its function is unknown (Lamb et al., 2007). Importantly, these variations in chemical environment surrounding the Schiff base can determine whether the linkage is protonated or unprotonated, affecting spectral sensitivity of the bound chromophore (Yokoyama et al., 2008, Hunt et al., 2004). With pigments that have unprotonated chromophore, their absorption maxima are in the UV range, as found in SWS1 cone opsins and RGR (Kono, 2006, Hao and Fong, 1996). With pigments that have protonated retinyl Schiff base, such as rod opsin, their peak absorbance values are red-shifted to the visible light region (Sakmar et al., 1989). The NPxxY$(x)_{5,6}$F motif (302-313) that connects TM7 and CL8 is another common feature found in the opsin family, where it provides structural constraints to stabilise the protein as the protein changes configuration during activation (Fritze et al., 2003). Once activated, the opsins must be deactivated to avoid continuous signalling. This is regulated by the binding of arrestin protein, which displaces downstream signalling components from the opsin after the opsin is phosphorylated by kinases during activation (Dolph et al., 1993). Interaction between arrestin with the opsin also prevents further activation by other G-proteins, desensitising the pigment molecule (Wilson and Applebury, 1993). Although the binding sites for arrestin is not strictly conserved, structural studies have shown that H8 and phosphorylated sites in the region play a crucial role in the affinity for arrestin (Kirchberg et al., 2011). In addition, a conserved Glu134 within an E(D)RY motif (134-136) located within TM3 has been shown to provide the negative charge required for stabilising the inactive opsin (Franke et al., 1992).

Although all opsins shared the above mentioned characteristics, different members of each subfamily have distinct molecular properties and functions, which are described below.

1.3.4 Vertebrate visual opsins

Classical vertebrate vision is mediated by visual pigments composed of retinal chromophores and two different types of visual opsins- cone-opsin (opn1) and rod-opsin (opn2 or RH1). Cone opsins are responsible for daylight or bright light vision and confer images with chromatic information, and are predominantly found in cone photoreceptor cells, whilst rod opsins are employed for night or dim light vision and are expressed in rod photoreceptors (Shichida and Imai, 1998).

1.3.4.1 Cone opsin (opn1)

The cone opsins can be divided further into four paralogous subgroups based on their amino acid sequences and spectral sensitivity: (1) long-wavelength-sensitive (LWS; perceived as yellow/orange/red) opsin with λ_{max} 510-560 nm, (2) short-wavelength-sensitive type 1 (SWS1; perceived as UV/violet) opsin with λ_{max} 360-430 nm, (3) short-wavelength-sensitive type 2 (SWS2; perceived as blue) opsin with λ_{max} 440-460 nm, and (4) middle-wavelength-sensitive (MWS or RH2; perceived as blue/green) opsin with λ_{max} 470-510 nm (Yokoyama, 2000, Davies, 2011). All four cone-opsin genes are present in most vertebrates, including ancestral lampreys (Collin et al., 2003b). However, it was found that primates lack the RH2 opsin but the MWS opsin is maintained. It has been postulated that RH2 opsin was lost as a result of the nocturnal bottleneck that occurred during the course of mammalian evolution

(Hunt et al., 2001), however, it was suggested that mesopic conditions were possibly more favourable given the order of gene loss and complement of genes still found in mammals compared to reptiles (Davies et al., 2012a). Old World monkeys, and apes have regained the *MWS* opsin by a duplication event that occurred in the *LWS* opsin gene, which resulted in green- as well as red-sensitive variants of the LWS pigment (Jacobs et al., 1996, Hunt et al., 1998). It is these three different cone opsins (red, green and violet-sensitive) which provide the basis for trichromatic vision in humans and primates. Interestingly, a similar duplication occurred in the elephant shark (Davies, et al., 2009), although, cone opsin duplications (and losses) are relatively common in fishes, especially the teleosts (Davies et al., 2012c). Genetic mutations in the *LWS* and *MWS* genes are associated with red-green colour blindness in humans. Studies have shown that red-green colour blindness is commonly caused by a polymorphism Ser180Ala in the LWS opsin (Asenjo et al., 1994), or gene deletion or generation of a chimeric LWS/MWS gene resulted from unequal recombination during meiosis (Deeb, 2004, Nathans et al., 1986). These changes either lead to the absence of LWS or MWS cone opsins, or cause the formation of abnormal cone pigments with altered spectral sensitivities. Deficiency in blue-yellow colour vision also occurs (known as tritanopia), and this disorder has been linked to three point mutations in the SWS1 opsin gene that leads to amino acid substitutions G79R, S214P, and P264S (Weitz et al., 1992a, Weitz et al., 1992b).

1.3.4.2 Rod-opsin (opn2)

Rod opsin (opn2 or RH1) is found in all vertebrate groups from lamprey to mammals. Evolutionary, it is most closely related to the RH2 cone opsin, followed by SWS2, SWS1, and then LWS opsins (Yokoyama, 1997). However, over 540 million years this opsin has evolved into distinct photopigment that is only expressed in rod photoreceptors. The rod opsin pigments have a λ_{max} of approximately 500 nm in blue-green light region (Hunt et al., 2001) and are extremely light sensitive, thus they play an important role in dim light vision. As an adaptation for increased visual sensitivity in low light environment, some deep-sea teleost fish, e.g. short fin pearleye (*Scopelarchus analis*) and mature Japanese eel (*Anguilla japonica*), express a second copy of the rod opsin gene that forms a pigment with λ_{max} blue-shifted by 40 nm and 6 nm respectively (Pointer et al., 2007, Zhang et al., 2000). Although another copy of *RH1* gene (*RH1-2*) is also found in two different species of zebrafish (*Danio roseus* and *Danio albolineatus*), its functional significance is unknown (Morrow et al., 2011). Studies have shown that in humans, mutations in *RH1* gene can cause retinal pathologies such as retinitis pigmentosa type 4 (RP4) (Dryja et al., 1990b, Farrar et al., 1992). RP4 is a blinding disease which involves progressive degeneration of retinal photoreceptor cells. Several mutations in the *RH1* gene associated with RP4 have been identified, amongst which are Pro23His, Met39Arg, Arg69His, and Pro53Arg substitutions (Davies et al., 2012b, Dryja et al., 1990a, Dryja et al., 1990b). The mutant RH1 proteins tend to impair cellular trafficking (Zhu et al., 2006) and/or aggregate with a regulatory protein called ubiquitin (Illing et al., 2002), disrupting their normal route to be targeted for ubiquitin-proteosome degradation. This results in

accumulation of mutant proteins in the photoreceptor cells, which leads to apoptosis (Saliba et al., 2002). Other mutations in RH1 (e.g. Gly90Asp, Thr94Ile, and Ala292Glu) have been associated with congenital stationary night blindness autosomal dominant type 1 (CSNBAD1), which is non-progressive impairment in night vision (al-Jandal et al., 1999, Dryja et al., 1993, Sieving et al., 2001).

1.3.5 Non-visual opsins

In vertebrates and invertebrates, there is a large group of non-visual opsins that are involved in extraocular photoreception with non-image forming functions (Foster and Hankins, 2002). These non-visual opsins possess important characteristics of a visual photopigment, including the lysine retinal attachment site, the presence of a glutamate counterion at site 113 or 181 and two conserved cysteine residues that form a disulphide bridge (as summarized in Fig. 1.7), but each is unique at both gene and protein level.

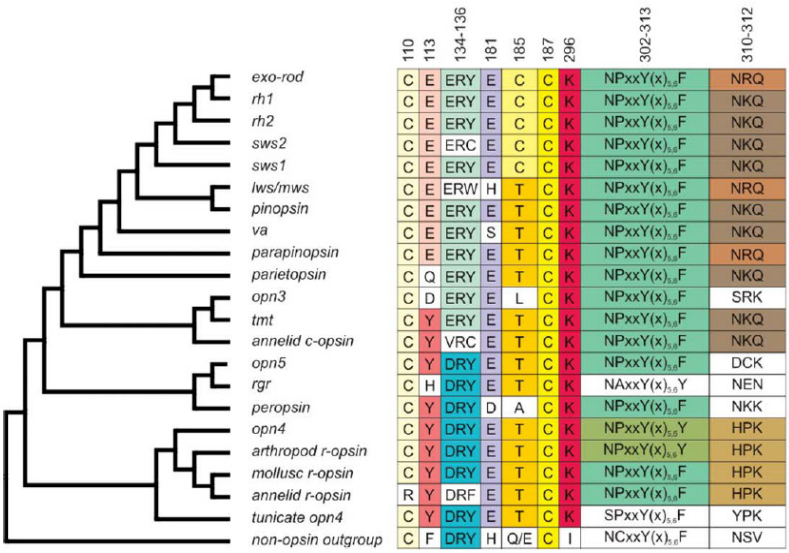


Figure 1.7 A simplified phylogenetic tree of all visual and non-visual opsin classes paired with a summary table of their conserved structural and functional characteristics (Davies et al., 2010).

1.3.5.1 Pinopsin (P-opsin)

Pinopsin (P-opsin) was the first non-visual opsin to be discovered by Okano and colleagues from chicken pineal organs (Okano et al., 1994). This opsin shares 43-48% amino acid identity with the vertebrate visual opsins. Experimental data have shown that *in vitro* reconstitution of P-opsin with 11-*cis* retinal generates a blue-sensitive photopigment that has a λ_{max} of ~462 nm (Max et al., 1998) and ~470 nm (Okano et al., 1994). P-opsin has now been identified from other bird species, e.g. pigeon (Okano et al., 1997), as well as from Reptilia (Kawamura et al., 1997) and Amphibia classes (Yoshikawa et al., 1998). Using immunohistochemistry, P-opsin was localised to the anterior preoptic nucleus of the hypothalamus in toad (*Bufo japonicus*) (Yoshikawa et al. 1998) and uniquely, in the pineal as well as retina of the diurnal gecko (*Phelsuma madagascariensis*) (Taniguchi et al. 2001). So far, orthologues of P-opsin gene have not been discovered in the agnathans, cartilaginous or teleost fish and mammals.

1.3.5.2 Vertebrate ancient (VA) opsin

Vertebrate ancient (VA) opsin was originally cloned from degenerate PCR screens of the cDNA generated from Atlantic salmon (*Salmo salar*) retina (Soni and Foster, 1997). Since phylogenetic tree analysis suggests that VA opsin diverged from a common ancestor at the very beginning of vertebrate evolution, prior to opsin families including the vertebrate visual opsins and P-opsin, it was given the name '<u>V</u>ertebrate <u>A</u>ncient' opsin (Soni and Foster, 1997). However, with the growing number of non-visual opsins being uncovered, it has been shown that many of these

(including parapinopsin, parietopsin, panopsin, teleost-multiple tissue opsin, neuropsin, peropsins and RPE-retinal G-protein-coupled receptor) actually arose before the origin of the VA opsin (Davies et al., 2010).

Functional studies have reported that reconstituting teleost VA opsins *in vitro* with 11-*cis* retinal can generate photopigments of λ_{max} between 451 nm to 510 nm (Soni and Foster, 1997, Kojima et al., 2000, Soni et al., 1998). Notably, it was the finding of VA opsin expression by RNA *in situ* hybridisation within a subset of horizontal, amacrine and retinal ganglion cells that provided strong evidence for non-cone, non-rod photoreception in the eye (Soni et al., 1998). Later studies demonstrated that VA opsin was also expressed in the pineal and subependymal cells of the anterior hypothalamus in teleost brain (Philp et al., 2000c, Kojima et al., 2000). Collectively, these data provided powerful implications for non-image forming irradiance detection in the central brain as well as in the retina.

Isolation of VA opsin from different teleost species revealed that within this opsin subclass there exist many isoform variants, which differs in lengths of their C-termini. The C-terminus of salmon VA opsin is composed of 19 amino acids (Soni and Foster, 1997). Isoforms of VA opsin with extended carboxyl-terminus (denoted as VAL) have been found in the common carp (*Cyprinus carpio*; 85 amino acids in length), and in zebrafish (*Danio rerio*; 74 amino acids in length) (Moutsaki et al., 2000, Kojima et al., 2000). There are also isoforms with medium length of the carboxyl-tail (denoted as VAM), such as that identified from the smelt fish (*Plecoglossus altivelis*; 49 amino acids in length) (Minamoto and Shimizu, 2002). It has been suggested by Kojima et al. (2008) that shorter splice variants of VA opsins could have been the

result of intron retention during splicing, whilst the longer variants were generated by splicing at conserved splice sites. The C-terminal tail possesses multiple phosphorylation and palmitoylation sites, which are important for regulating the function of opsin molecule. Since this region is variable between the long and short variants of VA opsins, it is likely that there is functional significance between the different isoforms.

Recently, orthologues of VA opsin have been identified in many other vertebrate classes, including the agnathans (e.g. lamprey) (Yokoyama and Zhang, 1997), teleosts (e.g. Atlantic salmon) (Philp et al., 2000b), and birds (e.g. chicken) (Halford et al., 2009). However, the *VA* opsin gene has not been found in mammals, which suggest that this opsin was lost early on during the evolution of mammalian lineage (Davies et al., 2010). In chicken, full-length VAL and VAS isoforms have been cloned from the eye and hypothalamic region of the brain (Halford et al., 2009). The amino acid sequences of chicken VAL opsin show ~35-40% identity with the visual cone/rod opsins and 41% identity to the pinopsin (Davies et al., 2010). UV/Visual spectrophotometry demonstrated that chicken VAL and VAS both exhibit an absorption peak at 490 nm, which closely matches the reported λ_{max} of avian photoperiodic response at 492 nm (Davies et al., 2012d). Linking these findings to physiological functions, Davies *et al.* (2012d) suggested that VA opsin is likely to have a key role in day length detection in birds.

1.3.5.3 Parapinopsin (PP opsin)

Parapinopsin (PP opsin) was originally identified from the pineal complex of catfish (Blackshaw and Snyder, 1997), however, orthologues have now been discovered in the teleosts (e.g. zebrafish) (Shiraki et al., 2010), frogs (*Xenopus tropicalis*) and lampreys (*Lampetra japonica*) (Koyanagi et al., 2004). Its expression pattern is highly selective, localised to parapinealocytes and a small subset of pineal cells. The reported sequence of the catfish PP opsin shows ~40% amino acid identity with the visual opsin group, and ~48% identity with P-opsin (Bellingham and Foster, 2002). Spectrophotometric analysis demonstrated that, when reconstituted with 11-*cis*-retinal, lamprey PP opsin forms a bistable UV-sensitive (UVS) photopigment which has a λ_{max} of 370 nm (Koyanagi et al., 2004). Irradiating the PP opsin pigment with UV light causes the retinal chromophore to undergo *cis* to *trans* isomerisation, generating a stable middle-wavelength-sensitive photoproduct with λ_{max} at 515 nm. It has been suggested that under environmental light, this bistable nature of the pigment and its ability to photoregenerate by visible light help to maintain a high level of UV-sensitive photopigment inside the cells (Koyanagi et al., 2004). This property of bistablity can also help the opsin to switch between different isomers of retinal and regenerate with fast kinetics, possibly for detecting rapid changes in down-welling radiation that comes from above water.

1.3.5.4 Parietopsin

The parietal eye (also known as third eye) of lizards contain a photoreceptor that has been shown to have two opsins co-expressed within the same cell, a blue-

sensitive P-opsin and a newly identified green-sensitive opsin called parietopsin (Su et al., 2006, Qiu et al., 2005). Unlike cone and rod photoreceptors, which hyperpolarise in response to light, the parietal eye photoreceptor can display two antagonistic light signalling responses, a hyperpolarising pathway driven by P-opsin that is maximally sensitive to blue light, versus a depolarising pathway mediated by parietopsin that is maximally sensitive to green light (Solessio and Engbretson, 1993). The peak spectral sensitivity of lizard parietopsin was reported to be 520 nm (Sakai et al., 2012), which is significantly red-shifted as compared to the spectral peaks measured for pinopsin (460-470 nm) (Okano et al., 1994, Max et al., 1998). These spectral absorbencies of pinopsin and parietopsin directly match the 'blue' and 'green' sensitivity of the parietal-eye photoreceptor cells, supporting their roles in mediating the observed light responses. Using single cell patch-clamp recording, the depolarising light responses mediated by parietopsin were linked to a G_o-pathway (Su et al., 2006), which resembles the signalling cascade of some invertebrate photoreceptors. In contrast, P-opsin has been shown to couple with gustducin in the parietal eye, a G-protein that is closely related to transducin (G_t) in visual photoreceptors (Su et al., 2006). Su and colleagues suggested that the G_o-pathway might be more ancient compared to the gustducin pathway, as it can be traced back to coelomates, which is the common ancestor of vertebrate and invertebrate mollusc. In the parietal eyes of ruin lizard, the presence of both pathways driving opposing responses to different wavelengths of light may represent a key step towards the development of dichromacy in vision. Parietopsin has also been linked to other

biological functions, which include detecting spatial orientation of the sun and electromagnetic field (Foa et al., 2009, Nishimura et al., 2010).

1.3.5.5 Extraretinal rod-like (Exo-rod) opsins

Since the early 1980s, it has been known that teleost pineal and parapineal organs exhibit opsin immunoreactivity (Vigh-Teichmann et al., 1982, Vigh-Teichmann et al., 1983). Independently, isolation of opsin sequences from the pineal of zebrafish (*Danio rerio*), pufferfish (*Fugu rubripes*), Atlantic salmon (*Salmo salar*), and lamprey (*Lethenteron japonica*) confirmed the presence of PP opsin, parietopsin, as well as a rod-like opsin (Philp et al., 2000c, Mano et al., 1999, Blackshaw and Snyder, 1997, Sakai et al., 2012). Given that the rod-like opsin was found outside the retina and its sequence closely resembles a rod opsin, it was designated "extraretinal rod-like (exo-rod) opsin". Thus far, exo-rod opsin has only been identified within the teleost class and not in any other vertebrates. This opsin is 74% identical to the visual rod opsin of the same species, which suggests that the two lineages diverged early on during the evolution of teleosts (Philp et al., 2000c). Although exo-rod opsin and rod opsin share similar spectral sensitivity (λ_{max} 498-502 nm), the two opsins exhibit very different kinetics in thermal bleaching of chromophore, with exo-rod opsin displaying a faster rate of spontaneous activation compared to rod opsin at temperature above 45°C (Tarttelin et al., 2011). Interestingly, Tarttelin and colleagues demonstrated that exo-rod opsin also undergoes a more rapid rate of decay at a specific meta-stable photoproduct state (Meta II) during activation. It was suggested that this shortened lifetime at Meta II state of exo-rod opsin may accelerate the

kinetics for its recovery from bleaching, contributing to its heightened sensitivity to light. Although it has been speculated that this opsin may be involved in the photic input of circadian system in teleosts (Philp et al., 2000a), its precise physiological roles have not yet been unresolved.

1.3.5.6 Panopsin/Encephalopsin (opn3)

Panopsin (Opn3) was originally called encephalopsin because it was found to be highly localised in the brain and testis of human and mouse, with very weak expression in the heart, liver and kidney (Blackshaw and Snyder, 1999). However, follow-up studies revealed that *Opn3* was in fact widely expressed in all the tissues examined, including retina, placenta, lung, skeletal muscles and pancreas (Halford et al., 2001b, White et al., 2008). Hence, the name 'panopsin' was proposed instead. Opn3 shows low amino acid identity (~30%) when compared with the classical cone and rod opsins. The counterion at position equivalent to 113 in bovine RH1, usually occupied by a negatively charged glutamate in vertebrate opsins, has been replaced by another negatively charged aspartate residue in Opn3 (Halford et al., 2001b). Nonetheless, Opn3 may still form a functional photopigment, as aspartate is also present as the counterion of *Xenopus* violet (UVS/UV) opsin (Starace and Knox, 1998). Mutation studies from RH1 showed that substituting Glu113 with aspartate generates a pigment that has λ_{max} slightly red-shifted by 5 nm (Sakmar et al., 1989). This spectral shift is due to the shorter side chain of aspartate compared to glutamate, causing the counterion to be slightly further away from the protonated Schiff base linkage (Sakmar et al., 1989). Genetic linkage and association analyses

have indicated that polymorphisms in the *Opn3* gene located on human chromosome 1q are likely associated with atopic and non-atopic asthma (White et al., 2008). Expression studies of the gene at RNA and protein level showed that *Opn3* is highly expressed in tissues which are involved in asthma (e.g. bronchiolar epithelium) and the immune system (e.g. macrophages and dendritic cells) (White et al., 2008). Together with functional experiments studying the effect of Opn3 in regulating interleukin-2 (IL2) secretion from T-cells, it seems that this opsin has an important role in immunomodulation and the pathophysiology of asthma (White et al., 2008).

1.3.5.7 Melanopsin (opn4)

Melanopsin (opn4) was first discovered in photosensitive dermal melanophores of *Xenopus laevis* (Provencio et al., 1998), where it was found to be also expressed in the eye and brain. Within the retina, melanopsin was detected in non-cone, non-rod cells including horizontal cells, RPE, and the iris (Provencio et al., 1998). In the brain, this opsin was found in the ventral part of magnocellular preoptic nucleus and the suprachiasmatic nucleus (SCN), both of these regions having been previously implicated in deep brain photoreception (Foster et al., 1994). These expression patterns prompted researchers to search for mammalian orthologues of melanopsin, which could potentially have a role in mediating light-dependent regulation of circadian rhythms. Indeed, human and mice orthologues of melanopsin were subsequently identified (Provencio et al., 2000), and localised in a small percentage (~1-2.5%) of intrinsically photosensitive retinal ganglion cells (pRGCs) (Sollars et al., 2003, Hattar et al., 2002). Since its discovery, melanopsin have been

identified in a broad range of species including cartilaginous fishes (e.g. elephant shark (Davies et al., 2012; in press)), teleost fish (e.g. zebrafish (Bellingham et al., 2002, Davies et al., 2011)), reptiles (e.g. Italian wall lizard (Frigato et al., 2006)), birds (e.g. chicken (Tomonari et al., 2005)), as well as marsupials (e.g. fat-tailed dunnart (Pires et al., 2007)).

More recently, it was discovered that two different melanopsin genes exist in non-mammalian vertebrates including zebrafish (*Danio rerio*), *Xenopus* (*Xenopus laevis)* and chicken (*Gallus gallus domesticus)*: (1) *mammalian-like melanopsin* (*opn4m*) and (2) *xenopus-like melanopsin* (*opn4x*) (Bellingham et al., 2006). It was shown that both of these genes encode a photosensitive opsin pigment, which triggers a G-protein signalling response when stimulated by light in the presence of retinaldehyde chromophore (Bellingham et al., 2006). The *opn4x* gene seemed to be lost in the mammalian lineage, as no *opn4x* gene was detected from *in silico* analysis of the eutherian, marsupial and monotreme genomes (Pires et al., 2007, Davies et al., 2010). The *opn4m* gene, on the other hand, has been detected in most vertebrates. Isoform variants differing by the length of their C-termini have been identified for *opn4x* as well as *opn4m* genes in various species. Chicken (*Gallus* gallus) was found to possess multiple isoforms for *opn4x* and *opn4m* (Torii et al., 2007). Recent data showed that long (*opn4xl*) and short (*opn4xs*) isoforms of *opn4x* are present in elephant shark (*Callorhinchus milii*) (Davies et al., 2012c). Two different isoform variants of *Opn4m*, *Opn4mL* and *Opn4mS*, were also found in mice (Pires et al., 2009) and human (Davies et al., 2012e). Although most of these isoforms form fully functional photopigments, their expression patterns are differentially regulated

throughout development of the animal (Hughes et al., 2012b, Verra et al., 2011). In mice, it has been demonstrated that these specific patterns of isoform expression directly correlate with the functional maturation of M1 and M2 cell type of pRGCs in retina (Hughes et al., 2012b). Hence, the isoform variants of Opn4 may play a role in regulating functional specialisation of the different subtype of pRGCs. In humans, there is an association of a P10L mutation of Opn4 with Seasonal Affective Disorder (SAD) (Roecklein et al., 2009), which is a type of depression that usually occurs during winter.

Based on several lines of evidence, melanopsin was initially proposed as a candidate photopigment that underpins circadian regulation in the newly discovered inner retinal photoreceptors, specifically the pRGCs. An early study using fluorescent immunocytochemistry demonstrated that *opn4* mRNA is expressed in about 2.5% of RGCs in the rat retina, majority of which belongs to type III group and exhibits intrinsic photosensitivity (Hattar et al., 2002). Retrograde tracing has shown that a large number of neurons from these Opn4-expressing RGCs (74%) projects into SCN region of the rat brain (Gooley et al., 2001). Since the SCN region contains a central molecular clock that syncs with environmental light/dark cycle, the link was made between melanopsin-positive pRGCs and photoentrainment of the central circadian clock via the retinohypothalamic tract directly to the SCN clock (Provencio et al., 2000). The property of intrinsic photosensitivity in melanopsin-expressing RGCs was further supported by whole-cell recording, which was measured to have a peak absorbance at 480 nm (Berson et al., 2002). Signicantly, the light-induced responses recorded from electrophysiology closely resemble the behavioural action

spectrum from circadian entrainment of rodents (Berson et al., 2002, Takahashi et al., 1984). Using mice which have their *Opn4* locus knocked in by a *tau-LacZ* reporter gene, it was revealed that β-galactosidase-positive pRGCs also target other central sites involved in the detection of ambient illuminance, sleep regulation and circadian photoentrainment, e.g. the intergeniculate leaflet, the olivary pretectal nuclei, the ventral subparaventricular zone, and the ventrolateral preoptic area (Hattar et al., 2002). From these findings, melanopsin was therefore suggested to be associated with divergent circadian photoresponses.

In an attempt to identify the precise functional roles of pRGCs, mutant mice without melanopsin (*Opn4*$^{-/-}$) were generated by several lab groups. These studies collectively showed that pRGCs expressing melanopsin are involved in a broad range of irradiance detection tasks, e.g. circadian phase shifting (Panda et al., 2002), photoentrainment of circadian rhythms (Panda et al., 2003), pupillary light reflex responses (Lucas et al., 2001), and melatonin suppression (Lucas et al., 1999). Though, these photoreceptive responses were only attenuated and not abolished in the melanopsin-null mice, suggesting there are inputs from other cells contributing to circadian entrainment. It was later found that extreme deficits in such phenotypes were only observed after *Opn4*$^{-/-}$ mice were crossed with mutant mice lacking rods and cones (Hattar et al., 2003, Panda et al., 2003). These rod-less and cone-less mice can occur either naturally with mutation in the *retinal degeneration* gene (*rd*/*rd*), which causes absence of rods and secondary losses of cones later on (Foster et al., 1991), or with both rods and cones ablated genetically (*rd*/*rd cl*) (Freedman et al., 1999). Notably, the rod-less, cone-less mice are still capable in maintaining

photoentrainment functions indistinguishable from the normal sighted mice. However, mice without rods, cones, and null for melanopsin completely lost the ability in regulating their circadian rhythms (Hattar et al., 2003). Taken together, the data (as summarized in Table 1.1) clearly show that the visual system of rods and cones complement pRGCs in regulating circadian photoentrainment. This conclusion is further supported by recent studies demonstrating the retention of UV-driven pRGC repsonses in *Opn4*$^{-/-}$ mice, which includes pupillary light constriction, sleep-induction, and circadian phase-shifting (Allen et al., 2011, van Oosterhout et al., 2012). It was found that these UV-light responses are specifically mediated by SWS cones and can occur independently of melanopsin, highlighting the significance of their contribution to NIF photoreception (Allen et al., 2011).

Although melanopsin knock-out studies have confirmed that melanopsin is an essential component for mediating photosensitivity in pRGCs, there remained a possibility that it might function as a photoisomerase, i.e. regenerating chromophore for a different opsin instead of directly activating signal transduction itself (Foster and Bellingham, 2002). In order to test the hypothesis that melanopsin is an actual signalling GPCR, various laboratories have used heterologous expression of

Table 1.1 A summary of phenotypes observed from transgenic experiments and melanopsin ablation studies in mice.

Experiments	Normal retina	Non-functional rods and cones (*rd/rd cl* or *rd/rd*)	Melanopsin knock out (*Opn4-/-*)	No rods, cones and Opn4 (*rd/rd cl* , *Opn4-/-*)
Vision?	Yes	No	Yes	No
Circadian photoentrainment capability?	Yes	Yes	Partial	No
Implications	Normal function	Opn4 alone can provide the needed signal for photoentrainment	Rods and cones contribute to photoentrainment	Either rods and cones or Opn4 is needed for adjusting to light

either human or mouse melanopsin and non mammalian orthologues (e.g. zebrafish, *Danio rerio*; and elephant shark, *Callorhinchus milii*) to induce photosensitivity in normally light-insensitive cells, e.g. Neuro2A cells (Melyan et al., 2005, Davies et al., 2011, Davies et al., 2012c), HEK293 cells (Qiu et al., 2005) and *Xenopus* oocytes (Panda et al., 2005). The results from these studies demonstrated that melanopsin expression alone was capable of activating a light and retinal-dependent phototransduction cascade with 9-*cis* or 11-*cis* retinal, inducing a change in cellular membrane potential. Furthermore, it has been shown that melanopsin also functions as a bistable photopigment, with inherent photoisomerase activity similar to the invertebrate opsins (Melyan et al., 2005, Panda et al., 2005, Davies et al., 2011, Davies et al., 2012c). Upon binding all-*trans* retinal, melanopsin uses long-wavelength light to regenerate 11-*cis* retinal. It was suggested that this light-driven reversibility allows melanopsin to sustain a prolonged response to light stimulation in pRGCs, which is different to the transient response observed in rods and cones (Mure et al., 2009, Wong, 2012). Further evidence supporting this comes from a recent paper using multi-electrode recording to measure the firing activities of melanopsin-expression pRGCs in mice retina (Sexton et al., 2012). Sexton and colleagues demonstrated that these melanopsin-expressing pRGCs are able to recover from bleaching within minutes after light activation, whilst rods and cones did not show recovery under the same testing condition.

These heterologous expression studies have also revealed the spectral characteristics of the melanopsin photopigment and provided important clues to its signal transduction pathway. In the presence of 11-*cis* retinal chromophore, the peak

spectral sensitivity of murine melanopsin recorded in human HEK293-TRPC3 cells and *Xenopus* oocytes is 480 nm (Panda et al., 2005, Qiu et al., 2005). These values correspond to the absorption peaks obtained from electrophysiology, and the action spectrum data of melanopsin-expressing pRGCs in rats (484 nm), primates (482 nm) and zebrafish (470 nm and 484 nm) (Berson, 2003, Dacey et al., 2005, Davies et al., 2011). Using specific antibodies and pharmacological blockers targeting different G-proteins and their downstream signalling components, this has led to a better understanding in the phototransduction cascade of melanopsin. Nonetheless, it is important to note that the current model of melanopsin signalling pathway is still far from comprehensive, as the precise identity of many components involved (e.g. G_{α} subunit, G_{β} subunit, downstream effectors and regulatory proteins) have yet to be identified. Based on existing data, melanopsin seems to utilise a phototransduction cascade similar to that of invertebrate rhodopsins. Invertebrate phototransduction pathways, modelled from extensive studies in the fruit fly (*Drosophila melanogaster*), involves the binding of a G_q/G_{11}-type G-protein to a rhodopsin molecule, leading to activation of phospholipase C (PLC) and then gating of transient receptor potential (TRP) channels, resulting in membrane depolarization (Hardie and Raghu, 2001). In contrast, the signalling cascade employed by visual opsins involves activation of transducin from the G_i/G_o-type G-protein class, phosphodiesterase, closure of CNG channels, resulting in hyperpolarization of the cell. Both melanopsin-expressing pRGCs and melanopsin-transfected Neuro-2A cells exhibit depolarization upon 480 nm light stimulation, resembling the photoresponse mediated by invertebrate rhodopsins. Antibodies against G_q/G_{11}-type (but not G_i/G_o-type) G-proteins and PLC

inhibitor have shown to significantly diminish melanopsin light-dependent responses in these expression systems (Melyan et al., 2005, Panda et al., 2003, Qiu et al., 2005). Collectively, the data suggest that melanopsin binds G_q/G_{11}, which in turns activates PLC-β and triggers the classical phosphoinositide pathway, leading to the transient increase in intracellular Ca^{2+} observed in the heterologous expression studies (Qiu et al., 2005, Isoldi et al., 2005, Panda et al., 2005).

Despite the strong evidence for activation of G_q/G_{11} type G-protein by melanopsin, detailed comparison with invertebrate phototransduction pathway revealed that there are still many unanswered questions in the current model of melanopsin signalling cascade. One notable area which has yet to be elucidated is the types of protein kinases responsible for modulating melanopsin signalling activity and its desensitisation mechanism (Hughes et al., 2012a). PKA and PKC are the two main classes of protein kinases that have been known to mediate regulation of GPCR-signalling dependent on secondary messengers. Phosphorylation activity of PKA is catalysed by the presence of cyclic adenosine monophosphate (cAMP), whilst that of PKC is activated by a membrane-bound lipid molecule called diacylglycerol (DAG), (Alberts et al., 2002). Since intracellular cAMP is driven by adenylate cyclase (AC), and is typically associated with G_{s-} and G_i-signalling pathways, it seems unlikely that PKA is employed in melanopsin signalling. However, it is now known that certain isoforms of the AC can be directly stimulated by the binding of $G_{\beta\gamma}$ subunit that is released upon activation of the G-protein (Diel et al., 2006, Tang and Gilman, 1991). Therefore, even though there is a lack of evidence for PKA-cAMP regulation in the melanopsin cascade, its involvement remains a possibility.

The secondary messenger system that activates PKC (DAG) is closely associated with calcium (Ca^{2+}) signalling, which is a classical response of the G_q/G_{11} pathway (Takashima et al., 2006, Melchior and Frangos, 2012). It has been shown that in invertebrates, PKC activity is linked to light adaptation and desensitisation of the photoreceptor responses (Yau and Hardie, 2009, Gu et al., 2005). Light adaptation (also known as background adaptation) occurs when the photoreceptor responds more quickly to increased intensity of illumination. This adaptation process also reduces sensitivity of the transduction pathway, by lowering the amplitude of each light response (Fain et al., 2001). Interestingly, both light adaptation and desensitisation have been reported in melanopsin-containing pRGCs, suggesting the potential involvement of PKC (Gamlin et al., 2007, Zhu et al., 2007). Nonetheless, there has yet to be direct evidence showing how and when in the melanopsin signaling pathway PKC may mediate the observed characteristics of photoresponses. Interestingly, functional genomics have indicated the involvement of a unique isotype of PKC (PKCz) in the melanopsin cascade, one which does not require Ca^{2+} and DAG for activation (Hughes et al., 2012a, Peirson et al., 2007). It was reported that knock-out mice lacking *PKCz* gene shared strikingly similar phenotypes as the *Opn4*$^{-/-}$ mice, with reduced pupillary light reflex, attenuated phase-shift circadian rhythm in respond to light, decreased period-lengthening under the effect of constant dim light, and deficiency in SCN expression of light-induced clock genes (Peirson et al., 2007). However, due to absence of association with the secondary messengers of melanopsin cascade, the involvement of PKCz in this pathway is yet to be proven.

In summary, despite the advancement in the current model of melanopsin cascade, the above mentioned aspects are a small part of the unexplored areas in melanopsin signalling pathway. Further work is much required for elucidating the precise mechanism and regulatory pathways that link melanopsin activity with pRGC responses.

1.3.5.8 Retinal-G-protein-coupled receptor (RGR) opsin and retinochrome

Retinal-G-protein-coupled receptor (RGR) opsin was originally isolated from screening the bovine RPE cDNA library (Jiang et al., 1993) and is thought to play a role as a photoisomerase catalysing the conversion of all-*trans*-retinal to its 11-*cis* form. Its expression is restricted to the RPE and Müller glial cells of vertebrates (Pandey et al., 1994). Its predicted amino acid sequence possesses a highly conserved lysine residue for retinal chromophore attachment in all opsins, but its putative counterion site 113 is occupied by a positively charged histidine (Briggs and Spudich, 2006). Other opsins usually possess either a negatively charged residue (e.g. glutamate) or a neutral residue (e.g. tyrosine) at site 113 as their counterions. Interestingly, a homologue of RGR in molluscs (retinochrome) also possesses a positively charged residue of either a methionine or a histidine at site 113 (Terakita et al., 2000). It has been shown that RGR-opsin shares 20.4% identity with the visual opsins and 23.6% identity to the retinochrome (Hara and Hara, 1967). Site-directed mutagenesis in retinochrome demonstrated that substituting H113 does not alter its spectral absorbance, but changing the negatively charged glutamate residue at site 181 causes deprotonation of the Schiff-base, leading to a blue-shift in its λ_{max} (Terakita

et al., 2000). This therefore suggests Glu181 acts as a counterion in the retinochrome, and this counterion switch is likely to be the same in RGR. The absorption spectrum of RGR photopigments reconstituted with all-*trans*-retinal is in blue light region (λ_{max} at 469 nm) (Hao and Fong, 1996). The presence of RGR contributes to the normal functioning of retina as mutations in this gene have been associated with retinitis pigmentosa (Morimura et al., 1999). It has been speculated that this opsin functions as a photoisomerase in the RPE for converting all-*trans* to 11-*cis* retinal in a light dependent manner, which is different to the light-independent system of the visual cycle (Radu et al., 2008). However, the role of RGR is complicated as *rgr*$^{-/-}$ mice are still capable of light-dependent regeneration of 11-*cis* retinal (Maeda et al., 2003). This indicates that although RGR acts in the retinoid cycle, its role in chromophore regeneration is not essential.

1.3.5.9 Peropsin (RRH)

Peropsin, also known as RPE-derived rhodopsin homologue (RRH), was initially isolated from the RPE of human and mouse retina (Sun et al., 1997). Subsequently, an orthologue of peropsin was found in the eyes of the lancelet (amphioxus; *Branchiostoma belcheri*) (Koyanagi et al., 2002). Although this opsin shows only ~26% amino acid identity with other photosensory opsins, phylogenetic evidence suggests that peropsin is closely related to the photoisomerase group which includes retinochrome and RGR (Sun et al., 1997, Bellingham et al., 2003) and neuropsin (OPN5). Immunohistochemistry and RNA *in situ* hybridisation showed that peropsin expression is confined to the RPE, specifically localised to the microvilli

around the outer segment of photoreceptor cells. The phylogenetic position and tissue expression pattern of RRH, when taken together, resulting in the speculation that peropsin may function as a retinal isomerase like RGR, maintaining a concentration of 11-*cis*-retinal inside the photoreceptor cells (Bellingham et al., 2003, Koyanagi et al., 2002). In fact, it has already been shown that amphioxus homologues of peropsin has the ability to bind all-*trans*-retinal and converts it into 11-*cis* form (Koyanagi et al., 2002), further supporting its role in retinal recycling. Despite the close association of RRH to the retinal regeneration function of RPE, there is currently no known mutation in the *RRH* gene linked to retinitis pigmentosa (Ksantini et al., 2007, Rivolta et al., 2006)

1.3.5.10 Neuropsin (Opn5)

Neuropsin (*Opn5*) gene was first discovered in the mouse and human genomes using bioinformatics (Tarttelin et al., 2003). The predicted amino acid sequence of this opsin possesses all of the conserved features of a functional opsin, which includes the conserved lysine residue in TM7 for chromophore attachment, a DRY motif for G-protein linkage, two conserved cysteine residues for forming a disulphide bond, and a conserved NPxxY(x)$_{5,6}$F motif in TM7. However, Opn5 only shares 25-30% identity with any other known opsins, thus it comprises a distinctive subfamily. Reverse transcription polymerase chain reaction (RT-PCR) data indicated that Opn5 is expressed in the eye, brain and testis of mouse, and in the eye and brain of human (Tarttelin et al., 2003). Immunohistochemistry revealed that within the mouse retina, Opn5 is localised to the RGCs and a subset of horizontal and amacrine

cells (Kojima et al., 2011). To date, opn5 has also been identified in chicken (Yamashita et al., 2010). It was found that apart from mammalian-type opn5, there are also two *opn5-like* (*opn5L1* and *opn5L2*) genes present in the chicken genome (Tomonari et al., 2008). Expression studies by Yamashita et al. (2010) and Tomonari et al. (2008) indicated that opn5 might be important in developing chicken retina, in particularly for the differentiation of ganglion cells and amacrine cells. Until recently, functional characterization of mouse and human Opn5 demonstrated that they are both UV sensitive with λ_{max} at 380 nm when reconstituted with 11-*cis* retinal (Kojima et al., 2011). The resulting photoproduct absorbs blue-light with λ_{max} at 470 nm, and it can be reverted back to UV-sensitive state with orange light illumination. Hence, Opn5 is a bistable UV-absorbing photopigment. Another paper has previously reported similar findings with a chicken homologue of mammalian Opn5, which has peak sensitivity in the UV region (λ_{max} at ~360 nm) and also exhibits similar bistability (Yamashita et al., 2010). Interestingly, there is also a report suggesting that the opn5 isolated from quails (*Coturnix japonica*) has its peak sensitivity at 420 nm (Nakane et al., 2010). This peak absorbance was determined based on an action spectrum determined from light-induced membrane current recordings. Since opn5 is likely to function as a bistable pigment that can be activated by two different wavelengths, it is possible that the λ_{max} obtained by Nakane and colleagues is an averaged value resulted from a mixture of the opsin and its photoproducts. GTPγS binding assay studies showed that both the mammalian and the chicken Opn5 signal through activation of Gi-type G-protein pathway (Yamashita et al., 2010, Kojima et al., 2011). These findings are significant because so far Opn5 is the only UV-sensitive

opsin reported in bird, and it has been suggested to be involved in avian photoperiodism (Nakane et al., 2010).

1.4 The discovery of *teleost multiple tissue* (*tmt*) opsin

Teleost multiple tissue (tmt) opsin is one of many non-visual opsins identified outside the classical ocular photoreceptor system. It was first discovered in zebrafish *(Danio rerio)* and pufferfish (*Fugu rubripes)*, and suggested to be the photopigment that regulates the peripheral circadian clock observed in many zebrafish tissues (Moutsaki et al., 2003, Whitmore et al., 2000). This opsin has a wide expression profile in both neural and non-neural tissues, even those previously not known to be photoreceptive, such as brain, heart, liver and kidney. Significantly, it is also found in a clock-containing embryonic cell line of zebrafish (PAC2), but not in the clock-deficient cell line (ZF-13; Whitmore, unpublished). Hence, the expression pattern of tmt opsin appears to coincide with the distribution of peripheral clocks. Although this may be a compelling correlation, there is currently a lack of functional data that directly links tmt opsins to photoentrainment of the peripheral clock in zebrafish.

1.4.1 Genomic organization and evolutionary history of tmt opsin

Tmt opsin is closely related to encephalopsin/panopsin (opn3), which also shows multiple tissue expression, but opn3 was initially thought to be present only in the mammals (Blackshaw and Snyder, 1999, Halford et al., 2001a, White et al., 2008) and thus the orthologue of tmt in the mammalian lineage. However, tmt and opn3 are distinct sister gene lineages, although grouped under the same subfamily, sharing ~41% amino acid identity (Moutsaki et al., 2003). Both *tmt* opsin and *opn3*

share a genomic structure common to the vertebrate opsins, with three consistent intron locations (intron 1, 3 and 4) (see Fig. 1.8, Moutsaki et al., 2003). This conservation pattern suggests that *opn3* and *tmt* opsins were likely to have descended from the same ancestral gene that gave rise to both visual and non-visual opsins. In addition, the matching genomic structure between opn3/tmt and parapinopsin suggests that they are more closely related to each other than with other opsin groups, which is in line with their phylogenetic positioning as shown in Figure 1.7, implying possible functional similarities between the two classes.

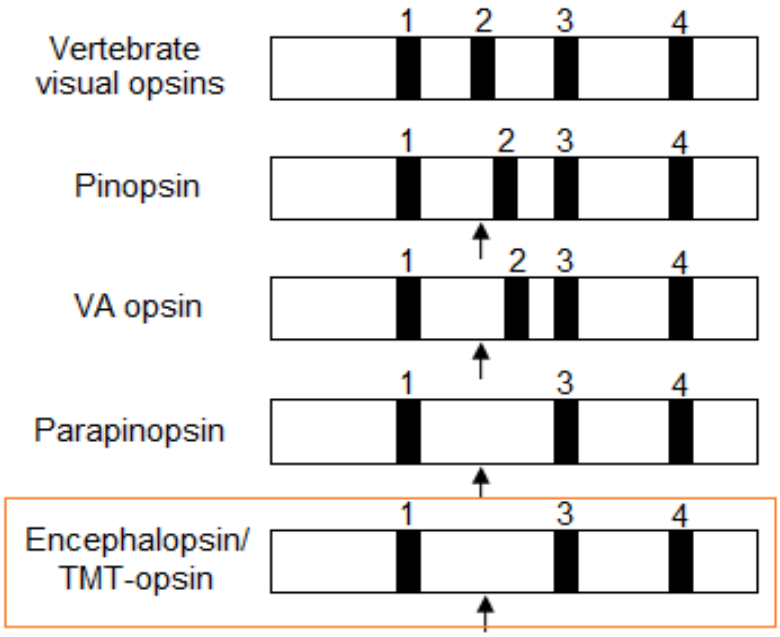


Figure 1.8 Genomic structures of vertebrate opsins and opn3/tmt-opsin. Exons in white boxes, and introns in black boxes (Moutsaki et al., 2003).

1.4.2 Structural characteristics of tmt opsins

Apart from possessing a gene structure showing high identity to other characterised photopigments, the published tmt opsin also contains conserved amino acid residues at positions that are known to be essential for opsin structural integrity and photosensory function. These include Cys110 and Cys187 for disulphide bond formation, a lysine residue at position 296 for a Schiff base linkage with chromophore, and a conserved E(D)RY motif (Moutsaki et al., 2003). Interestingly, the Schiff base counterion that is normally a glutamate in visual opsins is occupied by a tyrosine in pufferfish and zebrafish tmt opsins. However, this tyrosine substitution at the counterion position 113 is actually common in many vertebrate (non-visual) and invertebrate (visual) photopigments (Moutsaki et al., 2003). In fact, it was thought that the role of tyrosine as a counterion at this position predates that of the glutamate (Terakita et al., 2000). It is possible that, like many other non-visual opsins, the counterion of tmt opsin is displaced to Glu181 (Terakita et al., 2004).

1.4.3 Biological clock and circadian rhythms in animals

In most species, circadian rhythms are entrained to the environmental cues, such as light and temperature. These exogenous stimuli which can synchronise the internal clocks of an organism are termed as “zeitgeber”, which means “time giver” in German. Light is the most dominant zeitgeber for all circadian systems. It involves synchronising the oscillation of a core set of clock genes with the daily change in solar irradiance at dawn and dusk. Dawn and dusk serve as key stimuli for entrainment of circadian clocks because the rate of change in luminescence is most significant during these light-to-dark and dark-to-light transition periods.

Clock genes that are known to be involved in circadian regulation include, but are not limited to: (1) *Brain and muscle aryl hydrocarbon receptor nuclear translocator (ARNT)-like 1 (BMAL1)*; (2) *Circadian Locomotor Output Cycles Kaput (CLOCK)*; (3) *Period genes (PER1, PER2)*; (4) *Crytochrome genes (CRY1, CRY2)*; and (5) *Casein kinase 1 (CK1*). These genes encode for proteins which participate in autoregulatory transcriptional-translational feedback loops and exhibit intrinsic circadian rhythmicity (see Fig. 1.9). Photoentrainment of the rhythmicity is mediated by non-cone, non-rod photoreceptors and the photopigments expressed within them (Bellingham and Foster, 2002). In mammals, photoreception is confined to the eyes, with light inputs received by melanopsin-containing RGCs. These inputs are modulated by the cones and rods before being sent to a clock region in the brain called suprachiasmatic nuclei (SCN) (Freedman et al., 1999, Hattar et al., 2002). In non-mammals, however, the photoreceptive input and the circadian clock are multiple and can both take place in ocular as well as extraocular sites such as brain and pineal (Foster and Soni, 1998).

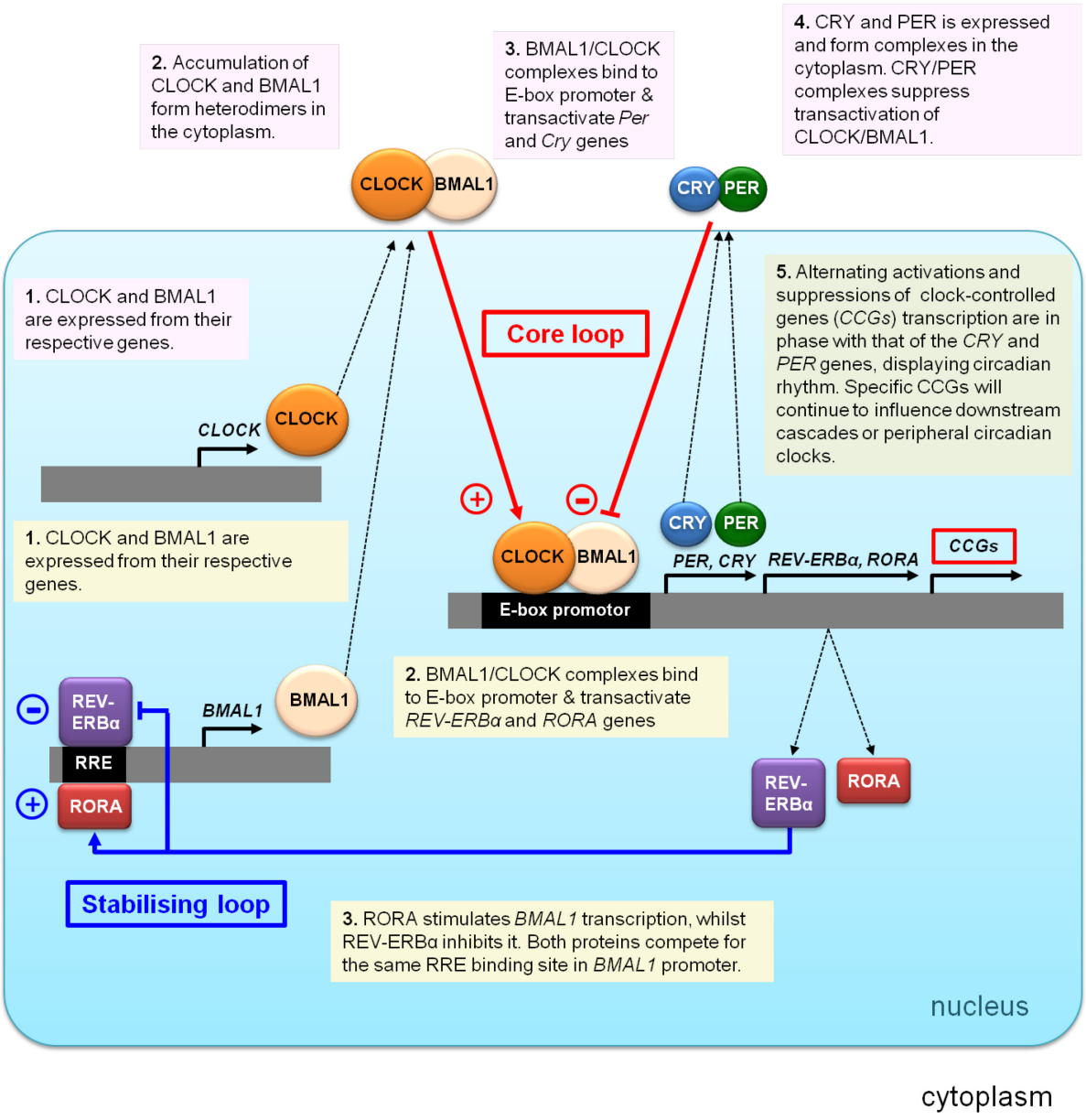

2. Accumulation of CLOCK and BMAL1 form heterodimers in the cytoplasm.
3. BMAL1/CLOCK complexes bind to E-box promoter & transactivate Per and Cry genes
4. CRY and PER is expressed and form complexes in the cytoplasm. CRY/PER complexes suppress transactivation of CLOCK/BMAL1.
CLOCK
BMAL1
CRY
PER
1. CLOCK and BMAL1 are expressed from their respective genes.
Core loop
5. Alternating activations and suppressions of clock-controlled genes (CCGs) transcription are in phase with that of the CRY and PER genes, displaying circadian rhythm. Specific CCGs will continue to influence downstream cascades or peripheral circadian clocks.
CLOCK
CLOCK
1. CLOCK and BMAL1 are expressed from their respective genes.
CLOCK
BMAL1
CRY
PER
PER, CRY
REV-ERBα, RORA
CCGs
E-box promotor
REV-ERBα
BMAL1
BMAL1
2. BMAL1/CLOCK complexes bind to E-box promoter & transactivate REV-ERBα and RORA genes
RRE
RORA
REV-ERBα
RORA
Stabilising loop
3. RORA stimulates BMAL1 transcription, whilst REV-ERBα inhibits it. Both proteins compete for the same RRE binding site in BMAL1 promoter.
nucleus
cytoplasm

Figure 1.9 The two transcriptional-translational feedback loops which generate a circadian rhythm in a single neuron of mammalian SCN. BMAL1 and CLOCK are the positive components in the circuit, whilst PERIOD (PER) and CRYPTOCHROME (CRY) are the negative components. In the "core" feedback loop (in blue), BMAL/CLOCK complexes function as transcriptional activators, binding to a conserved element called "enhancer (E)-box" within the promoters of *PER and CRY* genes to drive the production of their respective mRNA copies. The consensus motif for the E-box is CANNTG (Hastings and Herzog, 2004). As the resultant PER and CRY proteins build up over time, they form complexes that inhibit BMAL1/CLOCK transactivation, thereby feeding backward to suppress *PER* and *CRY* transcription. The "stabilising" feedback loop (in red) also involves the BMAL1/CLOCK complexes binding to E-box element, but this time within the promoters of *REV-ERBα* and *RORA* for transcriptional activation of these genes. The two proteins expressed by these genes can attach to the same "receptor response elements (RRE) promoter" of the *BMAL1* gene, but have opposing effects on its transcription. This competitive binding drives a rhythmic *BMAL1* mRNA production, completing the second feedback loop. Altogether, the output of this molecular clock is a precise periodic expression of clock-controlled genes (CCGs). Certain CCGs, such as *arylalkylamine N-acetyltransferase* (*aanat*), produces enzymes that are involved with the synthesis of hormones (e.g. melatonin), which can in turn modulate other nuclei or peripheral circadian clocks (Tosini and Menaker, 1996). Diagram drawn based on explanations by Shearman et al., 2000.

Traditionally, understanding vertebrate circadian organisation was based on the assumption that autonomous circadian oscillators reside solely within the CNS, i.e. the SCN of the hypothalamus in mammals (Prosser and Gillette, 1989) or the pineal in non-mammals (Collin et al., 1989, Takahashi et al., 1989), regulating overt rhythms. However, discovery of periodic clock gene expression in peripheral tissues has significantly extended the circadian clock model to a whole organism level. One of the first experiments demonstrating this was using immortalised cell-lines of rat fibroblast and hepatocyte cells, which were capable of generating circadian oscillation when stimulated by 50% horse serum (Balsalobre et al., 1998). To date, a wide range of peripheral tissues such as liver (Stokkan et al., 2001), heart (Cugini et al., 1993), and adipose cells (Zvonic et al., 2006) have also been shown to exhibit circadian rhythmicity of gene expression, although, the oscillation patterns observed in these isolated organs were only transient *in vivo*, as the rhythmicity diminishes after two to seven cycles, whereas those of the SCN were self-sustaining for a prolonged period (over 30 cycles) (Yamazaki et al., 2000, Yoo et al., 2004). By comparing the circadian phases in SCN-lesioned animals with the intact controls, the findings of Yoo et al. (2004) suggest that the SCN acts as a synchroniser rather than a pacemaker, as it functions to sustain the rhythm of peripheral clocks and maintain their "phases". The phase of a circadian rhythm is defined by the timing location of an identifiable reference marker during a complete cycle, such as the lowest body temperature, the beginning of sleep, or the onset of melatonin synthesis (Czeisler and Gooley, 2007). A shift in the circadian phase can therefore be detected when there is a change in the timing of the chosen reference marker from one cycle to another.

In "lower" vertebrates like the teleosts, it has been known that they possess an even more wildly distributed clock system. Many of the organs in zebrafish (*Danio rerio*) and reef fish (*Siganus guttatus*), including those outside the CNS (e.g. heart, kidney and spleen) exhibit strong circadian oscillation in clock gene expression (Whitmore et al., 1998, Park et al., 2007). The rhythmicity can persist for many cycles without diminishing even when the organs are isolated from the body. More intriguingly, these peripheral clocks appear to be directly entrainable by light (Whitmore et al., 2000, Pando et al., 2001). With the capabilities of phase-setting to external light and generating circadian oscillations in each peripheral tissue, they must independently contain a set of light detection machineries as well as the signalling cascade for mediating clock synchronicity. Interestingly, various cell-lines derived from zebrafish embryos, e.g. PAC2 (fibroblast cells) and Z3 (fibroblast-like cells), have also been shown to possess the properties of light detection and direct photoentrainment (Whitmore et al., 2000, Hirayama et al., 2005, Pando et al., 2001). These cell lines, unlike mammalian cell-lines that have been studied, did not require serum exposure to trigger clock function (Balsalobre et al., 1998, Whitmore et al., 2000). These observations suggest the presence of a photoreception system that controls the transcription of peripheral circadian clocks in teleosts.

1.4.4 A role for tmt opsin in regulation of peripheral clocks in fish?

The hypothesis that most organs, and in fact tissues of zebrafish, contain irradiance detectors was further supported by the identification of tmt opsin in teleost fish (Moutsaki et al., 2003). Since tmt opsin possesses all the principle features of a light-sensitive pigment at gene and protein sequence levels, and it is widely

expressed in various tissues of zebrafish, it has been proposed as a strong candidate for the entrainment of their peripheral clocks (Moutsaki et al., 2003).

1.4.4.1 Zebrafish: a model organism from a bright light environment

Zebrafish, *Danio rerio*, is a freshwater teleost fish that belongs to the family Cyprinidae. Its morphology is characterised by five pigmented, horizontal stripes along both sides of its body, which extend to the tip of its caudal fin (see Fig. 1.10A). Its anal fin also has the unique strip pattern. Adult zebrafish generally inhabit pelagic zone of the aquatic environment, such as streams, ponds and rivers (Spence et al., 2008). They are therefore exposed to the usual 24 h day/night cycles.

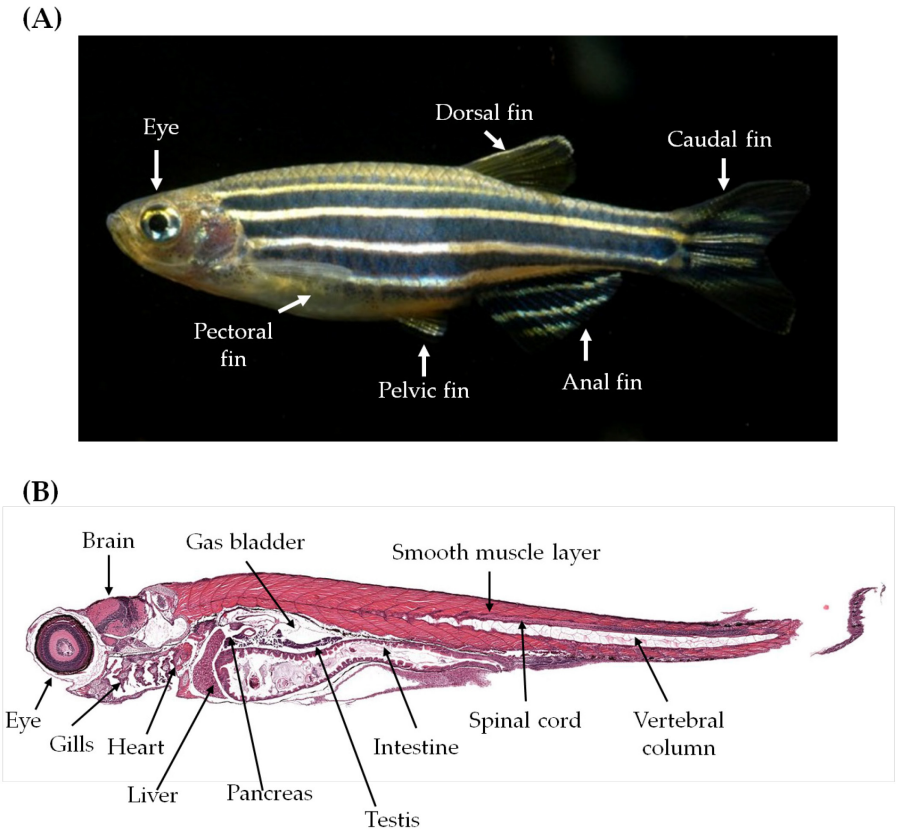


Figure 1.10 (A) Adult male zebrafish with external anatomy labelled (Noren, 2003) **(B)** A sagittal section of an adult male zebrafish (7 day post fertilisation; dpf) with internal anatomy labelled (picture was obtained from Zebrafish Atlas website and organs were labelled following the description in Menke et al., 2011).

The current prominence of zebrafish as a model organism for the study of clock regulation stems from the discovery of its highly decentralised clock system, a feature that is observed even in their embryonic cell-lines (Whitmore et al., 2000). Zebrafish circadian clocks seem to share the same set of clock genes as the mammalian system, involving: *bmal*, *clock*, *per*, and *cry* genes (Vatine et al., 2011). However, these clock genes are present in multiple copies in the zebrafish, with three *bmal* genes (*bmal1a*, *1b* and *2*), three *clock* genes (*clock 1a*, *1b* and *2*), four *per* genes (*per1a, 1b, 2* and 3), and six *cry* genes (*cry1a*, *1b*, *2a*, *2b*, *3* and 4) identified so far (Kobayashi et al., 2000, Wang, 2008a, Wang, 2008b, Wang, 2009). These copies of the clock genes suggest a complex regulatory mechanism that can fine-tune and maintain robustness of the circadian system. Similar to the mammalian clock, expression of most zebrafish clock genes is regulated by two transcriptional-translational feedback loops- a core loop and a stabilising loop (see Fig. 1.11). However, it has been demonstrated two of the zebrafish clock genes, *cry1a* and *per2*, are directly inducible by light (Tamai et al., 2007, Vatine et al., 2009). Light triggers the expression of a D-box binding protein called thyrotrope embryonic factor (TEF). TEF interacts with D-box element in the promoter region of *cry1a* and *per2* genes, which in turns drives transcription of the respective genes (Vatine et al., 2009). Hence, this explains why most peripheral tissues in zebrafish contain photoentrainable clocks. Nonetheless, the precise identity of the photoreceptor that controls the expression of TEF is still unknown. Thus far, tmt opsin has been suggested as the most likely candidate for mediating this peripheral photoreception function in teleost.

There are many advantages in using zebrafish as a model organism for the study of photobiology. As zebrafish embryos are translucent and accessible, the progression of their tissue development can be easily visualised and traced from early stages. The rate of embryo development also occurs very rapidly, with precursors of the major organs appearing 36 hours post fertilisation (hpf), and many of which (including the eyes) become fully formed 72 hpf. Figure 1.10B shows the internal organs in a 7-day old adult zebrafish from a sagittal section. In addition, zebrafish are amenable to a broad range of genetic manipulations, such as gene knock-down, mutagenesis, and transgenesis, allowing the generation of gene-loss and gain-of-function models for physiological studies. All these characteristics make zebrafish one of the most suitable models for characterizing the putative sensory mediator (tmt opsin) that may be responsible for the observed broad photoreception and its associated physiological functions.

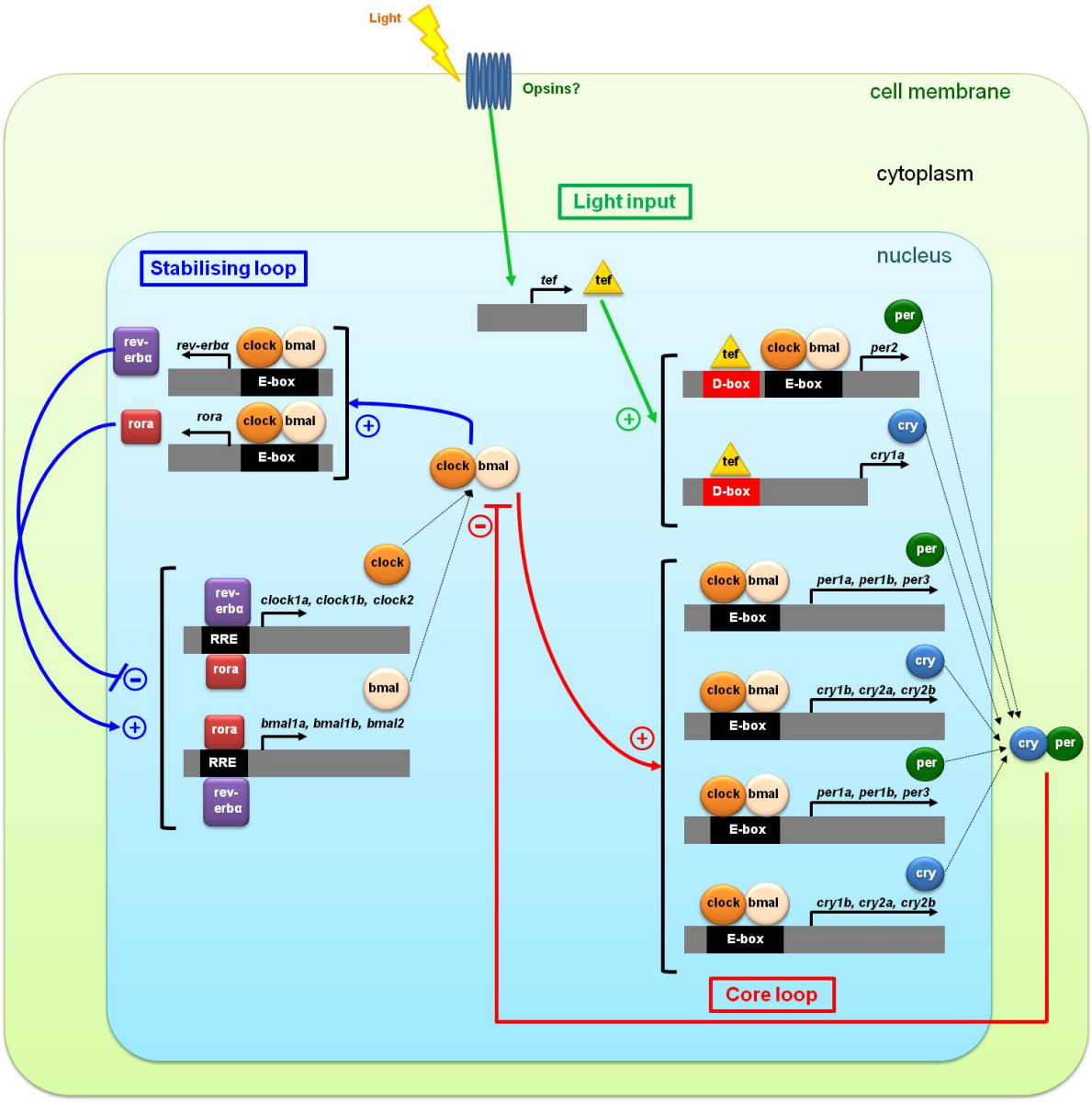

Light
Opsins?
cell membrane
cytoplasm
Light input
nucleus
Stabilising loop
tef
tef
rev-erbα
rev-erbα
clock
bmal
E-box
rora
rora
clock
bmal
E-box
clock
bmal
tef
clock
bmal
per2
D-box
E-box
per
cry
tef
cry1a
D-box
clock
rev-erbα
clock1a, clock1b, clock2
RRE
rora
bmal
rora
bmal1a, bmal1b, bmal2
RRE
rev-erbα
clock
bmal
per1a, per1b, per3
E-box
per
cry
clock
bmal
cry1b, cry2a, cry2b
E-box
cry
per
clock
bmal
per1a, per1b, per3
E-box
per
cry
clock
bmal
cry1b, cry2a, cry2b
E-box
cry
per
Core loop

Figure 1.11 Transcriptional and translational regulations of circadian clock genes in zebrafish. Two feedback loops similar to those in the mammalian circadian system are involved in the generation of rhythms: a "core" loop (in red) and a "stabilising" loop (in blue) (see Fig. 1.9 for details). In the core loop, expression of clock and bmal proteins forms heterodimeric complexes that can regulate the expression of *cry* and *per* genes. In the stabilising loop, clock/bmal complexes activate the transcription of *rev-erbα* and *rora* genes, which have opposing effects on *clock* and *bmal* gene expression. Unlike the mammalian system, two specific clock genes in zebrafish (*cry1a* and *per2*) are controlled by a transcription factor called thyrotrope embryonic factor (TEF) (Vatine et al., 2009). TEF specifically binds to a D-box element in the promoters of *cry1a* and *per2* genes, modulating their transcription processes. Significantly, the production of TEF themselves is directly inducible by light. Hence, clock gene expression in zebrafish is photoentrainable. Although the photoreceptor which collects light input to regulate TEF activity has not yet been identified, tmt opsin has been implicated as a promising candidate. Diagram is drawn based on explanation by (Vatine et al., 2009)

1.4.4.2 Mexican tetra: a model organism from a light-restricted environment

Mexican tetra, *Astyanax mexicanus*, is another freshwater teleost fish from the family Characidae. The existence of eyed surface-dwelling (surface fish, Fig. 1.12A) and eyeless or reduced eyed cave-dwelling (cavefish, Fig. 1.12B) forms in *Astyanax* made it a useful model for direct comparison with zebrafish and for the study of evolution in photoreception. The surface fish population generally lives within the epipelagic zones of the water (from surface to ~200 m in depth). Approximately 30 different cave populations of Mexican tetra have been recorded, mostly in the north-eastern regions of Mexico (Mitchell et al., 1977). The cave-dwelling forms, after being isolated in deep caves for 2.2 – 5.2 million years (Porter et al., 2007), have evolved troglomorphic adaptations, notably eye loss and depigmentation during their early stage of development (Culver and Pipan, 2009). Although small optic primordia are initially observed during cavefish embryogenesis, apoptosis of the lens induces their developmental arrest and causes the eyes to degenerate (Jeffery, 2005). These blind cavefish can, however, navigate in the dark using their lateral lines and enhanced mechanosensory system of cranial neuromasts to detect water movements (Jeffery, 2001).

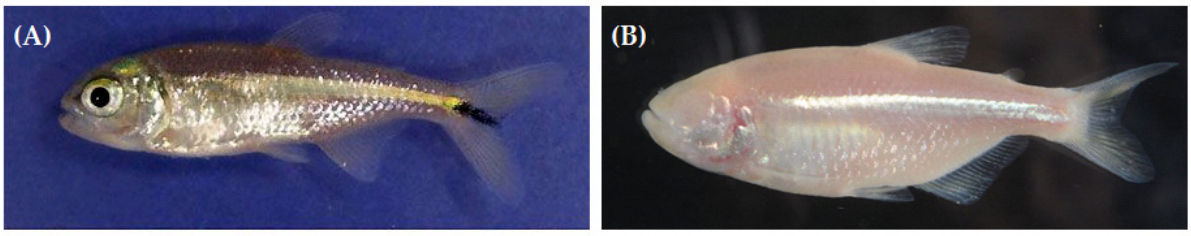


Figure 1.12 Mexican tetra **(A)** Sighted surface form (Yoshizawa, 2008) **(B)** Blind cavefish form (Borowsky, 2008)

Whilst evolution and development of the Mexican tetra have been studied extensively, many aspects of their circadian biology are still unexplored. In such extreme dark environments like deep caves, it is not surprising to find that the cave-dwelling forms no longer follow external 24 h day/night cycles (Cavallari et al., 2011). Nonetheless, the loss of eyes, and presumably the loss of visual image-formation, does not mean that they have completely lost the ability to detect light. In fact, some cavefish have retained photoentrainable clocks in physiological functions such as retinomotor activities (Espinasa and Jeffery, 2006). Recently, it has been reported that pseudogenes of *tmt* opsin and *melanopsin* (*opn4m2*) were identified in Somalian cavefish, *Phreatichthys andruzzii* (Cavallari et al., 2011). Both of these genes were found to have a frame-shift mutation that causes early termination in the translation process, forming truncated proteins. Significantly, the TM7 region that is missing in both mutant opsins contains a conserved binding site for retinal chromophore, which is required for photosensitivity. Cavallari et al. postulated that this loss in retinal-binding function by the two specific opsins has led to desynchronisation of the *P. andruzzi* clock. Subsequent experiment introducing full-length zebrafish tmt opsin and melanopsin into a cavefish cell-line (derived from the fin tissue of *P. andruzzi*) demonstrated that rescue of the photoentrainable clock is possible. Hence, these data gave a hint that tmt opsin and melanopsin may have photosensory roles in the regulation of cavefish peripheral clocks. Since the single point mutations observed in *P. andruzzi* melanopsin and tmt opsin could have occurred simply by chance, it would be fascinating to see if the same opsins in another cavefish species (*Astyanax mexicanus*) have remained intact through million years of constant darkness?

1.5 Research aims

The expression profile of zebrafish tmt opsin, together with its genetic and protein structure indicative of a functional pigment, as well as its ability to rescue light-entrainable circadian rhythm in *P. andruzzi* cell-line, have made it the most likely candidate as a photoreceptor potentially linked to peripheral clock entrainment (Moutsaki et al., 2003). In the present project, a combination of molecular biology, expression analysis and functional assays were used to elucidate the physiological roles of this opsin protein. This research thesis therefore aims to address the following questions:

(1) Are there any other tmt opsin(s) in the zebrafish genome?

- Hypothesis: Based on the strong evidence for whole genome duplication in the teleost lineage (Taylor et al., 2001, Taylor et al., 2003), it is likely that there are multiple genes encoding for *tmt* opsin in the zebrafish genome.
- For characterisation of the complete tmt opsin genomic repertoire in zebrafish, reverse transcription polymerase chain reaction (RT-PCR) and bioinformatics were used.

(2) What is the developmental and tissue expression profile of tmt opsins in zebrafish?

- Hypothesis: The expression pattern of different tmt opsins may be regulated developmentally as well as spatially.
- Expressions of tmt opsins in adult zebrafish tissues (eye and brain) and embryos (Day 1-5) were examined using NanoString nCounter analysis system and RNA *in situ* hybridisation (R-ISH).

(3) Do zebrafish tmt opsins respond to light, and if so, what are their peak spectral absorbencies?

- Hypothesis: With the presence of a Lys296 in the published tmt opsin (Moutsaki et al., 2003), this is indicative of its ability to bind retinal chromophore, and hence light sensitivity of a tmt pigment.
- Analysis was performed using UV-vis spectrophotometric analysis on recombinant tmt opsins that were expressed and reconstituted *in vitro* from Human Embryonic Kidney 293T (HEK-293T) cell lines.

(4) What are the spectral tuning mechanisms of tmt opsins contributing to the spectral sensitivity observed?

- Hypothesis: Based on initial phylogenetic data (fig. 1.7), it was deduced that the spectral tuning mechanisms of tmt opsins may be more closely related to the non-visual opsins (e.g. opn5 and opn4) than to visual opsins.
- Known counterions and spectral tuning sites were tested using site-directed mutagenesis.

(5) Can tmt opsins activate endogenous phototransduction by G protein signalling? If so, via which pathway?

- Hypothesis: As a member of the GPCR superfamily, characterised by seven TM segments, tmt opsins were predicted to bind specific G-protein ligand that could be coupled to a G protein signalling pathway.
- Functional assays including Ca^{2+} imaging, single-cell patch clamp electrophysiology, and luciferase-based bioluminescent assay were used in either murine neuroblastoma (Neuro-2A) or HEK-293T cells models.

(6) Are homologues of tmt opsins present in the cave and surface forms of Mexican tetra (*A. mexicanus*)? And if so, how do the structures of these opsins differ compared to those in the zebrafish?

- Hypothesis: There is a possibility that tmt opsins might have a functional role completely unrelated to light detection, regulating other aspects of physiology or developmental processes in teleost. Thus, these opsins could be maintained in the cave form of *A. mexicanus*, even though they have been living in complete darkness for millions of years.
- Comparison of the tmt opsins were made between cave and surface populations of *A. mexicanus*, as well as with zebrafish. This allows for the study of direct and adaptive effects of light on circadian photoentrainment.

CHAPTER 2

General Materials and Methods

CHAPTER 2: General Materials and Methods

2.1 Chemicals

2.1.1 QIAquick Gel Extraction Kit (Qiagen, UK)

- Buffer QG
- Isopropanol
- Buffer PE
- Buffer EB (10 mM Tris-Cl, pH 8.5)

2.1.2 GenElute Plasmid MiniPreps Kit (Sigma-Aldrich-Aldrich, UK)

- Resuspension solution
- RNase A solution
- Lysis solution
- Neutralization/Binding solution
- Column preparation solution
- Optional wash solution
- Wash solution concentrate
- Elution solution (10 mM Tris-HCl, 1 mM EDTA, pH approx 8.0)

2.1.3 HiSpeed Maxi kit (Qiagen, UK)

- RNase A solution (100 mg/ml; 7000 units/ml solution)
- LyseBlue
- Buffer P1- Resuspension buffer
- Buffer P2- Lysis buffer
- Buffer P3- Neutralization buffer
- Buffer QBT- Equilibration buffer
- Buffer QC- Wash buffer
- Buffer QF- Elution buffer
- Buffer TE (10 mM Tris-Cl, 1 mM EDTA, pH 8.0)

2.2 Animals

Zebrafish were raised and crossed following standard protocols in the laboratory of Professor Whitmore at University College London (UCL) (Dekens and Whitmore, 2008). They were kept in the aquaria at 29°C, fed twice daily and maintained under a 14 h day, 10 h night cycle. The fish facility has been approved by the Home Office in accordance with their regulations of animal treatment and care. All experiments involving the animals have been approved by the UCL ethical committee and were performed under animal licence number PIL 70/23714. The animals were sacrificed by rapid immersion in ice water and decapitation, which is in agreement with Schedule 1 of the Animal Welfare Act 2006 to ensure any suffering caused to be kept to a minimum. Mating tanks were prepared just before the start of night period, and embryos were collected the following five consecutive days (Day 0 to Day 5) at ZT3 (zeitgeber 3, i.e. 3 h after light onset).

2.3 Isolation and characterization of the zebrafish tmt opsin gene family

2.3.1 RNA source and preparation

Wild-type and albino adult zebrafish (4 months old) were sacrificed at ZT3-9. Dissections were performed in 1× phosphate-buffered saline (PBS; pH 7.0) using micro-dissection instruments under a dissecting microscope. Both the collected embryos and adult tissues (eye, brain, liver, gill, heart, caudal fins, muscles, skin, gut, testis and ovaries) were immediately processed for RNA extraction. Each of the tissues was homogenized separately in 1 mL of TRIzol reagent (Life Technologies,

UK) using micro-pestles to allow the release of RNA from the cells. All homogenized samples were stored at -80°C until required.

The homogenized samples were processed immediately for RNA isolation prior to use following manufacturer's protocol (Life Technologies, UK). Briefly, phase separation was performed on each sample by adding 0.2 mL chloroform and shook vigorously for 15 s by hand. The mixtures were incubated for 15 min on ice, followed by centrifugation at 15,000 rpm for 20 min at 4 °C. This separated the mixtures into a transparent upper aqueous phase containing RNA, a white interphase containing proteins, and a lower red phenol-chloroform phase. The aqueous phase of each sample was carefully transferred into new RNase-free eppendorf tubes. Total RNA was precipitated from the aqueous phase by addition of 0.5 mL 100% isopropanol, incubated on ice for 10 min, and then centrifuged at 18,000 x g for 20 min at 4 °C. Supernatant was removed from each tube, and the remaining RNA pellets were washed with 1 mL 75% ethanol. The samples were mixed briefly, and centrifuged again at 15,000 rpm for 10 min at 4 °C. After discarding the wash, the RNA pellets were air dried for 5-10 min and then resuspended in 50 μL of RNase-free water. All total RNA were also kept at -80°C for storage.

2.3.2 cDNA library synthesis and reverse transcription

Complementary DNA (cDNA) was synthesised *in vitro* from total RNA by reverse transcription, a process also known as first strand synthesis. The following components: 500 ng oligo$(dT)_{20}$ (Life Technologies, UK), 1-2 μg total RNA, and 20.5 μL water were combined in 0.2 mL thin-walled PCR tubes, incubated at 85°C for 15

min and chilled on ice for 2 min. Following this, 1× first-strand buffer, 10 μM DTT, 2.5 μL dNTP mixture (10 mM of each nucleotide), and 40 units (U) RNaseOUT Ribonuclease inhibitor (Life Technologies, UK) were added into each tube and incubated at room temperature for 2 min. 200 U of SuperScript III Reverse Transcriptase (Life Technologies, UK) was then put in. The reaction mixture was incubated at room temperature for 5 min, then at 50°C for 1 h before adding another 200 U of the reverse transcriptase and incubated for a further 1 h. At the final step, the reactions were warmed at 37 °C for 10 min and terminated at 55°C for 10 min. The cDNA made were stored at -20°C.

2.3.3 Reverse transcription polymerase chain reaction (RT-PCR)

The full-length coding region of zebrafish tmt opsins (tmt2l, tmt2s, tmt6, tmt9, tmt10, tmt14 and tmt24) were isolated by reverse transcription polymerase chain reaction (RT-PCR) from zebrafish adult (ZT3) retinal cDNA (prepared by Wayne I. L. Davies) and primers are listed in Table 2.1. Each reaction mixture contained 33 μl RNase-free water, 1× KOD DNA polymerase buffer, 1 mM $MgSO_4$, 0.2 mM dNTPs, 1 U KOD DNA polymerase (Novagen, UK), 40 ng of template, and 0.3 μM of each forward and reverse primer. PCR was conducted under standard thermal cycling conditions: an initial heating step at 95°C for 3 min, then through 40 cycles of 95°C for 30 s, 55°C for 30 s, and 70°C for 1 min 30 s, followed by a final cooling step at room temperature. The PCR products were ran on a 1.2 % (w/v) agarose gel by electrophoresis, excised and purified (as described in section 2.3.4-2.3.5 below). Even though all six tmt opsin PCR experiments gave amplicons of the correct size from

retinal cDNA, the concentration was only high enough to clone five out of the seven transcripts (tmt2l, tmt6, tmt9, tmt14, and tmt24) (Wayne I. L. Davies, personal communication), all of which were cloned into a mammalian expression vector (see section 2.5.1 below), and sequenced to ensure sequence fidelity prior to the start of this project by Wayne I. L. Davies. The remaining two tmt opsins (tmt2s and tmt10) were amplified from brain cDNA (prepared by Wayne I. L. Davies), where their higher expression level facilitated cloning in a mammalian expression vector via *Eco*RI and *Sal*I restriction enzymes. Opsin full-length amplicons were typically around 1 kb and ligated and subsequently transformed in *E. coli* (see section 2.5.1). Six colonies were picked for each experiment, and the inserts were sequenced in both orientations using vector-specific primers (Source BioScience, Oxford).

Table 2.1 *Oligonucleotides used to generate full-length zebrafish tmt opsin constructs*

Gene Specificity	Primers*	Oligonucleotide sequences**
ZF tmt2	PE_CHR2F	5′-GCGCGAATTCCACCATGATTGTGTCCAACTTGAGTG-3′
	PE_CHR2R1	5′-CGGCGTCGACGCTCCGTTATAATGGGCCACGAG-3′
	PE_CHR2R2	5′-CGGCGTCGACGCAAATCCAGCTATATCTTTCC-3′
ZF tmt6	PE_CHR6F	5′-GCGCGAATTCCACCATGTTTCCTGAAGAAACTAATATG-3′
	PE_CHR6R1	5′-CGGCGTCGACGCGCCAGAGACAGGGGTGCTGCAATCG-3′
	PE_CHR6R2	5′-CGGCGTCGACGCGTGCTCATTAAAAAGTTGGAAATG-3′
ZF tmt9	PE_CHR9F	5′-GCGCGAATTCCACCATGTTTTTCGAGCAGGCCGAT-3′
	PE_CHR9R1	5′-CGGCGTCGACGCTGGAATGACCTCTGGAGGGTT-3′
	PE_CHR9R2	5′-CGGCGTCGACGCTATATGACTTGTGAGTATGATATG-3′
ZF tmt10	PE_CHR10F	5′-GCGCGAATTCCACCATGGTCACTGTCCCCTTCCTCC-3′
	PE_CHR10R1	5′-CGGCGTCGACGCTTTGGGGGTGTCCATCAGGCT-3′
	PE_CHR10R2	5′-CGGCGTCGACGCTAAAGCACCTGTAAAACTGG-3′
ZF tmt14	PE_CHR14F	5′-GCGCGAATTCCACCATGGTCGTCTACATCTGGAGT-3′
	PE_CHR14R1	5′-CGGCGTCGACGCGGGTGTGTAGTGAACCACCAAGG-3′
	PE_CHR14R2	5′-CGGCGTCGACGCAAAACATATTGGTGCACCAAC-3′
ZF tmt24	PE_CHR24F	5′-GCGCGAATTCCACCATGATTGTGTCCAACTTGAGTG-3′
	PE_CHR24R1	5′-CGGCGTCGACGCTCCGTTATAATGGGCCACGAG-3′
	PE_CHR24R2	5′-CGGCGTCGACGCAAATCCAGCTATATCTTTCC-3′
*ZF EF1A****	QPCR_ZF_EF1A_F	5′-GCGGTACTACTCTTCTTGATGCCCTTGATG-3′
	QPCR_ZF_EF1A_R	5′-AGCCTCCATGGGTGGGTCGTTCTTGCTGTC-3′

Primers were designed by Wayne I. L. Davies.

* F = forward primer
R= reverse primer for amplifying the putative long isoform
R2 = reverse primers for amplifying the putative short isoform

** *Eco*RI site = GAATTC (red), *Sal* I site GTCGAC (blue)

*** Zebrafish *EF1A* was included as internal control for this experiment (Davies et al., 2011)

2.3.4 Agarose gel electrophoresis for separation of DNA fragments

1× tris-acetate-EDTA (TAE) buffer was made up from 50× stock (2 M tris-acetate, 0.05 M EDTA, pH 8.3) (Qiagen, UK). 1.2% (w/v) Hi-Res Standard Agarose (AGTC Bioproducts, UK) was dissolved in either 100 mL (small gels) or 200 mL (large gels) of the 1× TAE buffer by heating in microwave for 2 min. 0.5 μg/mL ethidium bromide (Sigma-Aldrich, UK) was added to the gel and mixed well before letting it set at room temperature. Once the gel had set, it was placed inside a gel tank of appropriate size and DNA samples were run in the same buffer at approximately 12.5 V/cm.

Depending on the expected size of DNA fragments, two different molecular weight markers were used to identify the size of samples: HighRanger Plus 100 bp DNA Ladder (Norgen Biotek, Canada; 0.5 μg per lane in each experiment) and Lambda DNA/HindIII Marker, 2 (Thermo Scientific, UK; 0.5 μg per lane in each experiment).

2.3.5 Gel extraction of DNA

After gel electrophoresis, DNA fragments were visualised by ethidium bromide staining in the gel on TM-20 UV transilluminator (UVP, Upland) and excised using a scalpel. Gel slices were purified using QIAquick Gel Extraction Kit (Qiagen, UK) following the manufacturer's protocol.

2.3.6 Molecular phylogenetic analysis of zebrafish tmt opsins

The nucleotide and amino acid sequences encoding seven transmembrane domains of zebrafish tmt opsins (tmt2s, tmt2l, tmt6, tmt9, tmt10, tmt14, and tmt24)

were aligned by codon-based matching using ClustalW software (Higgins et al., 1996) with default parameters. The alignments were refined manually for comparison with tmt opsins of other vertebrates from teleosts to mammals and with all other published visual and non-visual opsin sequences of the zebrafish. In addition, the cytoplasmic loops (CL 1, 2, and 3) and the C-terminal tail of zebrafish tmt opsins were also aligned separately with those of other known zebrafish visual and non-visual opsin sequences in the same manner.

Phylogenetic analysis was conducted on the codon-matched sequence alignments using Molecular Evolutionary Genetics Analysis (MEGA, version 4.0.2) program (Tamura et al., 2007). Neighbour-joining trees were generated by Wayne I. L. Davies based on two different methods: (1) Maximum Composite Likelihood (MCL) (Tamura and Nei, 1993) for nucleotide sequences, and (2) Jones-Taylor-Thornton (JTT) substitution matrix (Jones et al., 1992) for amino acid sequences, both with 1000 bootstrapping replications to measure the degree of support for internal branching. As an outgroup reference for the fish genes, the mRNA sequence of lancelet (*Branchiostoma belcheri*) opn4 (GenBank Accession Number: AB205400) was used. Opsin sequences that are included for construction of the trees are as follow: (1) extraretinal rod-like (exo-rod) opsin: zebrafish (*Danio rerio*), NM131212; (2) rod opsin (rh1): zebrafish (*Danio rerio*), NM131084; (3) rod opsin-like 2 (rh2): zebrafish (*Danio rerio*), NM131253 (rh2-1), NM182891 (rh2-2), NM182892 (rh2-3), NM131254 (rhr-4); (4) short-wavelength-sensitive 2 (sws2): zebrafish (*Danio rerio*), NM131192; (5) short-wavelength-sensitive 1 (sws1): zebrafish (*Danio rerio*), NM131319; (6) long-wavelength-sensitive and middle-wavelength-sensitive (lws/mws) opsins: zebrafish

(*Danio rerio*), NM131175 (lws-1), NM001002443 (lws-2); (g) vertebrate ancient (va) opsin: zebrafish (*Danio rerio*), AB035276 (va1), AY996588 (va2); (h) teleost multiple tissue (tmt) opsin: zebrafish (*Danio rerio*), tmt2s, tmt2l, tmt6, tmt9, tmt10, tmt14, tmt24 (GenBank Accession Numbers unassigned); (i) panopsin (opn3): zebrafish (*Danio rerio*), NM001111164; (j) neuropsin (opn5): zebrafish (*Danio rerio*), NM001200046; (k) retinal pigment epithelium-specific rhodopsin homolog (rrh) (peropsin): zebrafish (*Danio rerio*), NM001004654; (l) retinal G protein-coupled receptor (rgr): zebrafish (*Danio rerio*), NM001017877; (m) xenopus-like melanopsin (opn4x): zebrafish (*Danio rerio*), GQ925718 (opn4x-1), GQ925719 (opn4x-2); (n) mammalian-like melanopsin (opn4m): zebrafish (*Danio rerio*), GQ925715 (opn4m-1), GQ925716 (opn4m-2); GQ925717 (opn4m-3); (o) chordate melanopsin (opn4): lancelet (*Branchiostoma belcheri*), AB205400 (outgroup).

2.3.7 Bioinformatic tools

Several bioinformatic programs were used for analysing each tmt opsin sequence:

(i) **BLAST v2.2.26** (Mount, 2007)

URL: http://blast.ncbi.nlm.nih.gov/Blast.cgi

To search for similarities between the query sequence and a database of the published nucleotide or protein sequences, nucleotide blast (nucleotide) and tblastx (translated nucleotide) programs were used.

(ii) **CLUSTAL W2** (Larkin et al., 2007)

URL: http://www.ebi.ac.uk/Tools/msa/clustalw2/

This program was used to align two or more inferred amino acid sequences of the opsins. It helped to identify areas of similarities that may be more highly conserved than other regions in the sequences.

(iii) **TMHMM Server v. 2.0**

URL: http://www.cbs.dtu.dk/services/TMHMM-2.0/

This program was used for predicting transmembrane domains in proteins based on a "Hidden Markov Model" (HMM). Mapping the deduced amino acid sequence to the computational model, TMHMM captures information such as hydrophobicity, charge bias, and helix lengths to infer the membrane topology of the tmt opsins. The outputs were shown as a list of the location of amino acid regions being predicted as helices, cytoplasmic loops or extracellular loops. Probability plots were also generated to illustrate the likelihood for the predicted locations of the residues (see table 3.3).

(iv) **NetNGlyc 1.0 Server**

URL: http://www.cbs.dtu.dk/services/NetNGlyc/

For prediction of N-linked Glycosylation sites (Asn, N) in proteins that follows the consensus motif of Asn-X-Ser/Thr (N-X-S/T) (Bause and Legler, 1981).

(v) **GPS 2.1 for Phosphorylation**

URL: http://gps.biocuckoo.org/

For computational prediction of potential phosphorylation sites with serine and threonine kinases in the proteins.

(vii) **CSS-Palm**

URL: http://csspalm.biocuckoo.org/

For prediction of palmitoylation sites at the C-terminus.

2.4 Expression studies of tmt opsin in adult zebrafish and embryo tissues

2.4.1 Analysis of tmt opsin expression by RT-PCR

RT-PCRs were performed using cDNA synthesised from mRNA extracted from zebrafish embryos (Day 2 to 5) and adult tissues (brain, eye, testis, ovaries, heart, gut, gill, caudal fins, muscles, and liver). The primers used for the detection of tmt opsin cDNAs in these tissues were the same as those used in cloning full-length coding sequences (see Table 2.1). PCR mixtures for the amplification from cDNA were prepared in 25 μL, containing 16 μL sterile water, 0.5 μL KOD DNA polymerase, 1.5 mM $MgSO_4$, 2.5 μL 10× KOD DNA polymerase buffer, 0.2 mM dNTPs, 0.5 μL template cDNA, and 0.3 μM of each forward and reverse primer. PCRs were conducted as follows: an initial polymerase activation step at 95°C for 5 min, then 40 cycles of 95°C for 30 s, 55°C for 30 s, and 70°C for 1 min, followed by cooling at room temperature. The amplicons were electrophoresed in 1.2% agarose

gels containing ethidium bromide. Gels were photographed using BioDoc-It *UV Transilluminators* (UVP, Upland).

2.4.2 Analysis of tmt opsin expression by RNA i*n situ* hybridization (R-ISH)

2.4.2.1 Digoxgenin (DIG) probes design and synthesis

Constructs containing the full-length coding sequences for all six tmt opsins (tmt2, tmt6, tmt9, tmt10, tmt14, tmt24) (6 μg of each) were digested with *Eco*RI and *Sal*I restriction enzymes and subcloned into pBluescript II SK phagemids (Stratagene, UK). The pBluescript vector carries a promoter for T7 RNA polymerase to produce reverse complementary RNA. All clones were verified by sequencing before being used as templates for the synthesis of digoxgenin (DIG)-labelled antisense riboprobes by using the DIG RNA Labeling Kit (SP6/T7) (Roche, UK). Each tmt opsin-containing pBluescript plasmid was linearised by *Eco*RI enzyme prior to use. Following the manufacturer's protocol, run-off transcription reactions were set up to generate antisense riboprobes from the linearised plasmid containing a T7 promoter. DIG-UTPs were incorporated into the transcripts during the transcription process. The labelled probes were then purified using NucAway™ Spin Columns (Life Technologies, UK) to remove salt and unincorporated nucleotides. The probes were further purified by ethanol precipitation with 4M LiCl. Purified riboprobes was resuspended in 100 μL of *in situ* hybridisation buffer (Life Technologies, UK) and stored at -80°C.

2.4.2.2 In situ hybridisation on brain and retinal cryosections

Eyes and brains (from albino female zebrafish) were collected at ZT3 for cryosectioning. They were fixed in 4% paraformaldehyde (PFA) in 1× PBS (pH 7.0) at 4°C overnight. The eyes and brains were washed three times in 1× PBS and cryoprotected in 30% sucrose for at least 48 h before embedding in OCT medium (VWR, UK). These tissues were frozen at -80°C until they were sectioned. Slices of ocular (18 μm) and brain (10 μm) tissues were cut at -20°C (for eye) or -18°C (for brain) using a cryostat (Leica Biosystems, UK). The tissue slices were place on Superfrost Plus glass microscope slides (VWR, UK) and stored at -80°C until use.

Prior to experimentation, slides were air dried for 30 min and re-fixed in 4% PFA in 0.1 M phosphate buffer (PB) (0.081 M Na_2HPO_4, 0.019 M NaH_2PO_4; pH 7.4) for 10 min at room temperature. Tissues were treated with 1 μg/ml solution of Proteinase K (Sigma-Aldrich, UK) in 50 mM Tris-HCl (pH 7.5) and 6.25 mM EDTA (pH 8.0) for 5 min, followed by acetylation in acidified triethanolamine for 10 min at room temperature. Prehybridisation was performed on the tissue sections at 68°C for at least 1 h in ULTRAhyb® Ultrasensitive Hybridisation Buffer (Life Technologies, UK). The prehybridisation buffer on each slide was replaced with 100 ng antisense DIG-labelled riboprobes (from section 2.4.2.1) in 500 μL of fresh hybridisation buffer. Slides were carefully covered with clean coverslips and incubated at 68°C overnight in a Techne Hybridiser HB-1D oven. After 12-16 h, the cover slips were removed in 5x saline-sodium citrate (SSC) solution pre-warmed to 72°C and the slides washed twice in 0.2× SSC solution (Life Technologies, UK) (72°C with 40 min per wash). Slides were then equilibrated in B1 buffer (0.1 M Tris, pH7.5; 0.15 M NaCl) for 5 min

at room temperature, prior to the addition of blocking solution (10% heat-inactivated goat serum (HINGS) (Sigma-Aldrich, UK) in B1 buffer) and an incubation for 1 h at room temperature. The blocking solution was subsequently removed and the slides were incubated with a monoclonal anti-DIG antibody conjugated to alkaline phosphate (1:5000; Roche, UK) in B1 buffer with 1% HINGS at 4°C overnight. Detection of the hybridised probes was achieved by using the BCIP/NBT AP Substrate Kit IV (Vector Laboratories, UK) following the manufacturer's instructions. During the developing process, slides were kept in closed, humidity containers for 4 to 5 days before the reaction was terminated by immersion in water. Glycergel Mounting Medium (Dako, UK) was used to mount coverslips onto the slides.

All results were viewed with a 5× Plan-Neofluar lens (brain sections) or a 40X Neofluar lens (retinal sections) on a Zeiss Axioplan 2 Microscope (Carl Zeiss, USA) and photographed with a colour digital camera (Micropublisher, Canada). Tiling of acquired photographs was performed using Image-Pro Insight® version 8.0 software (Media Cybernetics, USA). Anatomical boundaries of the brain sections were defined using a topological atlas (Wullimann et al., 1996).

2.4.2.3 Whole-mount in situ hybridization on zebrafish embryos

Zebrafish embryos were harvested on days 1 to 5 at ZT3 from the fish facility at UCL. Embryos were developed in culture flasks, which were kept in a 28°C water bath. Unfertilized and unhealthy eggs were discarded each day using a Pasteur capillary pipette. 1 mL of 200 μM 1-phenyl 2-thiourea (PTU), a tyrosinase inhibitor, was added to the flasks 24 h post fertilisation to block pigmentation during

development and to enhance visualisation of internal structures. Embryos at the appropriate developmental stages were fixed in 4% PFA in 1× PBS overnight at 4°C in 1.5 mL eppendorf tubes. The next day, the chorions of day 1 to day 4 embryos were removed with a sharp forceps. Dechorionated embryos were washed three times in 1× PBS and dehydrated in 100% methanol for 15 min at room temperature. Embryos were stored at -20°C in 100% methanol for at least 2 h prior to use.

Whole-mount *in situ* hybridisation of the embryos was performed based on a published protocol (Thisse and Thisse, 2008). Rehydration and permeabilisation steps were carried out in 1.5 ml eppendorf tubes at room temperature on shakers. Embryos were rehydrated by rinsing in successive dilutions of methanol (MeOH) in PBT (1× PBS, 0.1% Tween-20): 5 min in 75% MeOH/PBT; 5 min in 50% MeOH/PBT; and 5 min in 25% MeOH/PBT. This was followed by two 5 min washes in 100% PBT. Day 2 to Day 5 (D2-D5) embryos were permeabilised with Proteinase K (10 μg/mL) in PBT, and the duration of this treatment varied depending on the developmental stage: 2 min for D2 embryos; 5 min for D3/D4 embryos; and 10 min for D5 embryos. Proteinase K digestion was terminated by washing the embryos twice in PBT for 5 min and re-fixing them in 4% PFA in 1× PBS for 20 min. A prehybridisation step was carried out by incubating each tube of embryos with 250 μL hybridisation mix (HM) (50 μg/mL heparin, 500 μg/mL RNase-free tRNA, 5× SSC solution, 50% deionized formamide, and 0.1 % Tween-20; pH 6.0 adjusted by citric acid) for 2 h at 65°C. HM was then discarded and replaced with 250 μL fresh HM containing 100 ng antisense DIG-labelled riboprobes (see section 2.4.2.1). Embryos were hybridised with riboprobes overnight at 65°C.

After hybridisation, embryos were subjected to successive washes at 65°C with gentle shaking: 5 min in HM; 30 min in 50% HM in 2× Saline-Sodium Citrate (SSC); and two 35 min washes in 0.2× SSC before cooling to room temperature. Additional rinses were performed at room temperature with shaking by washing the embryos in 1× PBS for 15 min, followed by two 5 min washes in PBS. Prior to antibody staining, embryos were incubated in 2% blocking agent (Roche, UK) in maleic acid buffer (MAB) for at least 3 h at room temperature. The blocking buffer was then replaced with anti-DIG antibody solution (diluted 1:5000) in 2% blocking agent in MAB and incubated at 4°C overnight. Immunocytochemical detection of the riboprobes was carried out using BM purple staining solution (Roche, UK). Embryos were washed four times, 30 min per wash, in 1× PBS at room temperature before being equilibrated three times, 5 min per equilibration, in 1 mL BM purple staining buffer. Embryos were then transferred into six-well plates containing 1 ml developing substrate and incubated in the dark at room temperature until the colour reaction was completed. The detection reaction was terminated by several washes in PBT, and then re-fixing in 4% PFA in 1× PBS for 2 h at room temperature. Embryos were rinsed in 1× PBS before mounting in 75% glycerol in 1× PBS and stored at 4°C.

2.4.3 Quantification of tmt opsin mRNA expression by NanoString nCounter

Expression levels of zebrafish tmt opsins were quantified by nCounter® Analysis System (NanoString Technologies, UK), which was carried out by UCL Genomics. A tissue panel of adult zebrafish eyes, brains, hearts, livers, PAC2 cells, Day 1 and Day 2 embryos were collected by Dr. Kathy Tamai (UCL) at ZT3 in TRIzol

Reagent (Life Technologies, UK), prior to the extraction of total RNA using RNeasy mini spin columns (Qiagen, UK). RNA concentrations were measured with a Nanodrop (Thermo Scientific, UK) prior to analysis. CodeSets comprising reporter and capture probe pairs specific for detecting tmt opsins were designed and constructed by UCL Genomics. Each capture probe carries a biotin label at the 3′-end for immobilisation of the target gene onto nCounter cartridge surface, whilst each reporter probe carries a unique colour code at the 5′-end for barcoding the gene of interest (Figure 2.1). The probe pairs contained target-specific sequences that can together recognise a 100 nucleotide region within the tmt opsin mRNA sequence (see Table 2.2).

Briefly, 5 μL of diluted total RNA (~100-150 ng) was hybridised with both capture and reporter probes at 65°C for 16 h. The hybridised samples were then loaded onto the nCounter Prep Station for automated post-hybridisation processing, including the removal of unhybridised probes, washes, and immobilisation of the target/probe tripartite complexes. A cartridge containing the loaded samples was then transferred to the nCounter Digital Analyser for imaging and quantification. For analysis of expression, the counts for all target genes in every sample were normalised based on positive spike-in controls to account for any variation between different lanes. Following this, a background count was determined from the negative controls by taking an average of the most stringent negative values. Triplicate experiments for each sample were averaged and recorded. A target gene was considered to be present if the average count of the samples was significantly greater than the background unit (i.e. greater than 15; $P<0.05$).

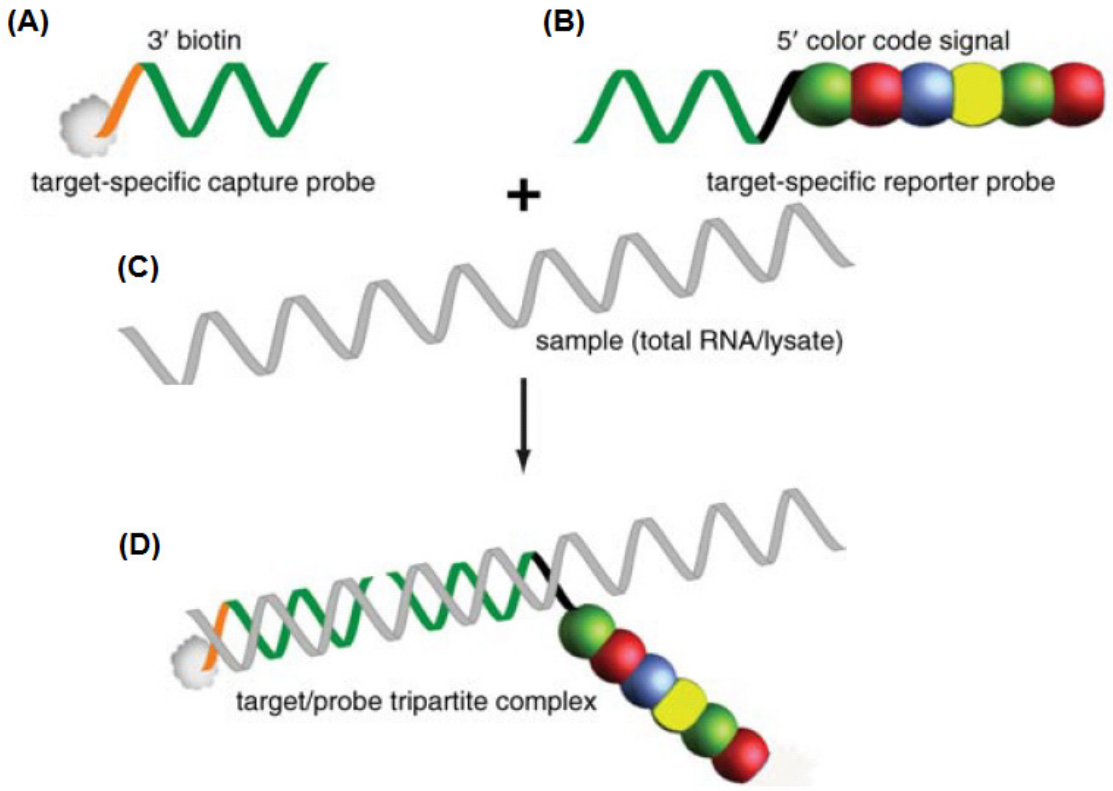


Figure 2.1 Diagram demonstrating the target-specific nature of **(A)** capture and **(B)** reporter probes (Kulkarni, 2011). The capture probe has a 3′-biotin handle, whilst the reporter probe carries a 5′-barcode. The orange region in the capture probe is a 30 base sequence common to all capture probes, and the black segment in the reporter probe is a 60 base sequence common to all reporter probes. These common regions were used for the removal of excess unbound probes post hybridisation washes. A successful hybridisation reaction with **(C)** target mRNA at 65°C would result in the formation of **(D)** target/probe tripartite complexes. The probe-labelled mRNAs are purified by binding to the surface of nCounter cartridge via a streptavidin-biotin linkage. It is the 5′-barcodes on the immobilised complexes that are read by the nCounter Digital Analyser, thus giving an absolute number of transcripts. Since there is a unique barcode for each probe, multiple probes can be incubated with the same RNA sample, giving a more accurate measure of relative expression levels.

Table 2.2 *Probe sets for zebrafish tmt opsins used in NanoString analysis*

Gene Specificity	Target region	Target sequence
*ZF tmt2** (common)	581-680	GGACGACCAAGACCGCCAATAACATCTCCTACATCATTTG CCTCTTTATTTTTTGCCTGATCGTCCCCTTCCTGGTCATCA TATTCTGCTATGGGAAGCT
ZF tmt2 long	956-1055	TTCTCAACTGCGACAAACCACAACGAGGTTCTAGTCTGAA AAGCTCTTCAAAGACCAAACCTTTTCGTCCCGGACGCCGC ACGGACAACTTCACTTTTAT
ZF tmt6	88-187	CCTGCTGAGAAACTGTCACGGACGGGACACAATGTAGTC GCTGTTATTTTGGGATCTATTCTAATATTCGGGACCCTGAA CAATCTTGTAGTCCTCGTTC
ZF tmt9	1025-1124	TGAACCGCAGGGTCAACAGTAACGCTGTAGCCTGCACCG CTCAGATCTCCACCGGGACTCACAACCATGATTGCAGCAC TCACGTCACAGAGAGGAGCAA
ZF tmt10	549-648	TTCAGTCCAGTGGAACAAGCGCTCTCCGGAAACCCGGTC TTACGTCATCTGCCTCTTTGTCTTCTGTCTTCTGCTGCCTC TCCTTCTGATGGTCTACTGC
ZF tmt14	256-355	ATCAGCGTCAGTGACATACTGGTCTGCTTGTTTGGGACTC CCTTTAGCTTCGCGTCCAGTCTCTATGGGAAATGGCTTTT GGGACATCACGGCTGCAAAT
ZF tmt24	131-230	GCCTTGGATTCATCGGAACTTTCGGCTTTCTCAACAACAC GCTCGTGCTCGTCCTGTTCTGTCGCTACAAAGTGCTGCGC TCGCCTATGAACTGTCTCCT
ZF RPL13A	397-496	GGCTGCTCTGGACAGGCTGAAGGTGTTTGATGGCATCCC TCCACCTTATGACAAGAGAAAGCGCATGGTTGTCCCAGCT GCTCTCAAGATTGTGCGTCTG
ZF EF1A	1350-1449	GTGCGTGACATGAGGCAGACCGTTGCTGTCGGCGTCATC AAGAGCGTTGAGAAGAAAATCGGTGGTGCTGGCAAGGTC ACAAAGTCTGCACAGAAGGCTG
ZF UBIQUITIN B	1577-1676	CTGTCCGACTACAACATCCAGAAAGAGTCCACCTTGCATC TGGTGTTGCGTCTCAGGGGAGGCATGCAGATTTTTGTTAA AACCTTGACAGGAAAGACCA

* *tmt2* (common) was designed to detect both *tmt2* long and short isoforms.

** Zebrafish *RPL13A, EF1A and ubiquitin B* were used as internal controls for this experiment.

2.4.4 5′-Rapid amplification of cDNA ends (5′-RACE)

5′-RACE was performed with ocular and brain mRNA obtained from adult zebrafish to isolate the 5′-end of tmt opsin transcripts, using a FirstChoice®RLM-RACE kit (Life Technologies, UK) according to the manufacturer's instructions.

Total RNA was reverse transcribed into cDNA as described in section 2.3.2. The initial RNA templates were degraded from the RNA-cDNA hybrids by adding 2 U of an endonuclease RNase H (New England BioLabs Inc., UK) and incubated at 37°C for 20 min. The first-strand cDNA was then purified by QIAquick PCR purification kit following the manufacturer's protocols (Qiagen, UK). For 5′-RACE analysis, two nested PCR reactions were set up to acquire the 5′-end of the opsin mRNA sequences from the first-strand cDNA synthesised. Sequences of forward (F1, first round PCR; F2, second round PCR) and reverse (R1, first round PCR; R2, second round PCR) primers are listed in table 2.3. Each 50 μL reaction in the first round of PCR comprised 2.5 U BIOTAQ™ DNA polymerase (Bioline, UK), 1× NH_4 buffer, 2 mM $MgCl_2$ solution, 10 μM dNTP mix, 50 ng reverse transcribed product, 0.2 μM of 5′ RACE outer forward (F1) and *tmt*-specific reverse outer primer (R1). This first round of PCR was run initially at 94°C for 5 min, followed by 40 cycles of 94°C for 30 s, 55°C for 1 m, 72°C for 1 m 30 s, and then a final extension step of 72°C for 5 min. The second round of PCR was set up in the same manner using 2 μL of the first-round product as template, but with the 5′-RACE inner forward primer (F2) and *tmt*-specific reverse inner primer (R2). Identical thermal cycling conditions were used as in the outer 5′-RACE PCR. Whole volumes of each reaction were run in 1.2% (w/v) agarose gels containing 1μg/mL ethidium bromide and visualised on a UV

transilluminator (UPV, Upland). Discreet PCR amplicons were identified, excised and purified as previously described.

Table 2.3 *Primer sequences used in 5′-RACE of zebrafish tmt opsins*

Nested primer name	Primer sequence (5′ to 3′)	Description
5′-RACE Adapter	GCUGAUGGCGAUGAAUGAACACUGCGUUUG CUGGCUUUGAUGAAA	Anneal to 5′-end of decapped mRNA
5′-RACE Outer primer	GCTGATGGCGATGAATGAACACTG	Nested 5′-RACE outer forward primers
5′-RACE Inner primer	CGCGGATCCGAACACTGCGTTTGCTGGCTTT GATG	Nested 5′-RACE inner forward primer
tmt2_5′-RACE_R1	CTCAAGCAGATGTTCACCAGCAGAAAGTTG	5′-RACE tmt2 exon 1 outer reverse primer
tmt2_5′-RACE_R2	CCCGATGAATCCCAAGCTCACCGCCACGAG	5′-RACE tmt2 exon 1 inner reverse primer
tmt6_5′-RACE_R1	GTCCGCAGTGTCTTGAATTTACAAAACAGA	5′-RACE tmt6 exon 1 outer reverse primer
tmt6_5′-RACE_R2	ACAAGATTGTTCAGGGTCCCGAATATTAGA	5′-RACE tmt6 exon 1 inner reverse primer
tmt9_5′-RACE_R1	CACACCAACATGTCACTGATGCTGATGTTA	5′-RACE tmt9 exon 1 outer reverse primer
tmt9_5′-RACE_R2	TGAATCCCAGGAACACAGAGAGCGCGATGA	5′-RACE tmt9 exon 1 inner reverse primer
tmt10_5′-RACE_R1	AGACAAACCAGCATATCGCTCACAATGATA	5′-RACE tmt10 exon 1 outer reverse primer
tmt10_5′-RACE_R2	AGACAAACCAGCATATCGCTCACAATGATA	5′-RACE tmt10 exon 1 inner reverse primer
tmt14_5′-RACE_R1	GGATCAGGTTTATAGGAGTCCACAGTGTGC	5′-RACE tmt14 exon 1 outer reverse primer
tmt14_5′-RACE_R2	GCAGACGGCGGTCACCGTGTGGCCGGTCCT	5′-RACE tmt14 exon 1 inner reverse primer
tmt24_5′-RACE_R1	GTAGCGACAGAACAGGACGAGCACGAGCGT	5′-RACE tmt24 exon 1 outer reverse primer
tmt24_5′-RACE_R2	GATGAATCCAAGGCACACCGCGACTACCAA	5′-RACE tmt24 exon 1 inner reverse primer

2.4.5 Cloning and sequencing of 5′-RACE PCR fragments

5′-RACE PCR fragments amplified by BIOTAQ™ DNA polymerase have an 'A' overhang at their 3′-ends. Using the Promega pGEM®-T East Vector System I kit, the 5′-RACE PCR products were cloned into a linearised pGEM®-T Easy vector (Promega, UK) which carries a 5′-terminal 'T' at both ends. Following the manufacturer's protocol, each ligation reaction contained 1× Rapid Ligation Buffer for T4 DNA Ligase, 0.025 pM pGEM®-T vector, 0.076 pM PCR products (insert), 1 μL T4 DNA ligase, and 2 μL nuclease-free water. The reactions were incubated overnight at 4°C before transformation into *E.coli* DH5*α* competent cells (Life Technologies, UK). Recombinant clones which carried the inserts were identified by blue/white screening on LB agar plates containing 100 μg/mL ampicillin, 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and 40 μg/mL X-gal. To ensure accuracy of the sequence data, twelve positive colonies (white) from independent 5′-RACE PCR reactions were picked for inoculating starter cultures of 2 mL LB medium, each containing 100 μg/mL ampicillin. The starter cultures were incubated at 37°C with shaking for 6 h. 500 μL of each starter culture was then used to inoculate 250 mL LB medium with ampicillin (100 μg/mL), which were grown at 37°C for 12-16 h with shaking for large-scale plasmid production. Plasmids were subsequently isolated by using a GenElute Plasmid MiniPreps Kit (Sigma-Aldrich, UK) according to the manufacturer's instructions and confirmed by test digestion with *EcoR*1 restriction enzyme. The selected positive clones were sent with M13 forward primer (5′-CACGACGTTGTAAAACGAC-3′) (Promega, UK) and amplicon-specific reverse primers to Source BioScience LifeScience (Oxford, UK) for sequencing. M13 forward

primer is designed to facilitate the sequencing of inserts which have been cloned into *lacZ*-containing vectors such as pGEM-T. It specifically anneals to position 2941- 2957 of the pGEM-T vector (Promega, UK). The identity of each tmt opsin cDNA sequence was confirmed by bioinformatics through an online nucleotide BLASTn program (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch) (Altschul et al., 1990).

2.4.6 Promoter analysis of zebrafish tmt opsins using bioinformatics

The position of the transcription start site of a gene correlates with the location of its promoter. Identifying major transcription start site for each zebrafish tmt opsin gene is therefore crucial for localising their promoter regions and for elucidating the basic mechanisms of transcriptional regulation. Transcription start site for each tmt opsin transcript were first determined from the 5′-flanking region on genome sequences using the Neural Network Promoter Prediction (NNPP) version 2.2 software (Berkeley Drosophila Genome Project, http://www.fruitfly.org/seq_tools/promoter.html, with 'eukaryote' as setting of the organism and a minimum cut-off score of 0.8). The flanking region of the putative start site was then compared to the consensus sequence of the initiator (INR) element, YYC**A**NWYY with '**A**' representing the transcription start site (Liston and Johnson, 1999). Once mapped, a region of 3 kb upstream to the transcription start site was subjected to a targeted search for regulatory factor binding motifs using a variety of online programs as described below.

The upstream region was first scanned for CpG islands (Cross and Bird, 1995) and CCCTC-binding factor (CTCF) insulator elements (Phillips and Corces, 2009), which are typically associated with transcription regulatory factors at or near the start of a gene. The search for CpG islands was achieved by http://zeus2.itb.cnr.it/cgi-bin/wwwcpg.pl, and the putative binding sites for CTCF were identified by direct comparison to the consensus sequence CCGCNNGGNGGCAG (with a criterion of >0.7 as cutoff score value) (Ishihara and Sasaki, 2002). Core and proximal promoter elements within 200 bp upstream of the TSS were then determined: (a) TATA box (consensus sequence TATAWAWN) (Juo et al., 1996) (http://zeus2.itb.cnr.it/cgi-bin/wwwHC_TATA.pl), (b) CCAAT box (consensus sequence GGNCAATCT) (Benoist et al., 1980), (c) GC box (consensus sequence KRGGCGKRRY) (Sogawa et al., 1993), (d) B recognition element (BRE; consensus sequence SSRCGCC) (Lagrange et al., 1998), (e) Motif ten element (MTE; consensus sequence CSARCSSAACG) (Lim et al., 2004), and (f) Downstream Promoter Element (DPE; consensus sequence RGWCGTG) (Burke et al., 1998). MATCH™ version 1.0 database (http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi) and NSITE database (http://linux1.softberry.com/berry.phtml?topic=nsite&group=programs&subgroup=promoter) were also used to search for general transcription factor binding sites (TFBS), such as Oct-1 binding site (consensus sequence ATGCAAAT) (Phillips and Luisi, 2000) and AP-1 binding sites (consensus sequence TGASTCA) (Lee et al., 1987), as well as those known to be involved in retinogenesis e.g. paired box gene 6 (pax-6) (Wawersik and Maas, 2000, Walther and Gruss, 1991) within a 3 kb upstream flanking region of the *tmt* opsin transcription start site.

In addition, the 3 kb upstream region containing the promoter of *tmt* opsins was investigated for putative transcription factor binding sites that are known to be tissue-specific. These binding sites were determined by manual comparison of the genomic sequences with published consensus binding sites (a cut-off score of >0.7 was applied). Those that are specifically important for photoreceptor development and functions include: nuclear receptor subfamily 2 group E member 3 (nr2e3) (consensus sequence RAGRTCAAARRTCA) (Chen et al., 2005); rar-related orphan receptor beta (ror$_{\beta}$) (consensus sequence WWAWBTAGGTCA) (Jetten et al., 2001), which is also involved in circadian clock gene regulation (Ueda et al., 2005); cone-rod homeobox (crx) (consensus sequence TAATCA) (Furukawa et al., 1997); retina-specific region-1/photoreceptor conserved element-1 (ret-1/pce-1) (consensus sequence CAATTAG) and orthodenticle homeobox/bovine AT-rich sequence-1 (otx/bat-1) (consensus sequence TGATTAG) (Kimura et al., 2000); retina-specific region-4 (ret-4) (consensus sequence GCTTAG) (Chen and Zack, 1996); neural retina leucine zipper (nrl) [consensus sequence TGAN$_{6-8}$GCA (Rehemtulla et al., 1996), where the N$_{6-8}$ core has similarities with the binding site for AP-1 (consensus sequence TGASTCA) or the 3′-5′-cyclic adenosine monophosphate (cAMP) response element (TGAGCTCA) (Kataoka et al., 1994, Kerppola and Curran, 1994)]; glass-like binding motif (consensus sequence ACCCTTGAAATGCC) (Moses and Rubin, 1991); and *Caenorhabditis elegans* homeobox-10 (ceh-10) containing homolog (chx10) (consensus sequence YTAATTRR) (Dorval et al., 2006). Binding sites that are specific for other tissues include: homeobox gene nkx2.5 in cardiac development and differentiation (Benson et al., 1999, Guner-Ataman et al., 2013); hepatocyte nuclear

factors (HNF-1, HNF-3β, HNF-4) in liver regeneration and function (Costa et al., 2003); and E26-like transcription factor 1 (Elk-1), which is involved in signalling pathways in the brain (Sgambato et al., 1998) and as well as lung and testis development (Rao et al., 1989). Additionally, known regulatory motifs associated with the circadian clock genes were also investigated based on manual comparison (with a minimum of 70% nucleotide matches). These include Enhancer-box (E-box; consensus sequence CANNTG) (Massari and Murre, 2000), cAMP response element (CRE; consensus sequence TGAGCTCA) (Lonze and Ginty, 2002), ROR elements (RREs; consensus sequence WAWNTRGGTCA) (Harding and Lazar, 1993) and Destruction-box (D-box; RTTAYGTAAY) (Falvey et al., 1996).

2.5 Functional characterisation of zebrafish tmt opsins

2.5.1 Generation of tmt-pMT4 constructs for protein expression

Amplicons of tmt opsins from RT-PCR in section 2.3.3 were cloned into pMT4 mammalian expression vectors to allow for expression of the tmt opsin proteins in cells. The primers used for amplification of tmt opsin cDNA were designed to incorporate an *Eco*RI restriction site at the 5′ end of the coding region, and a *Sal*I restriction site at the 3′ end to aid subcloning (see Table 2.1 for detail of primer sequences). The forward primer (F) for each gene contained a translation start codon (AUG) within a Kozak consensus sequence, which allowed for efficient translation of the recombinant protein. The reverse primers substituted the predicted stop codon in exon 4 (for putative long isoforms- R1 primers) or in the intron after exon 3 (for putative short isoforms- R2 primers) of the opsin protein with a *Sal*I restriction site,

which read into a sequence encoding for bovine rhodopsin ID4 (Rho 1D4) epitope tag (ETSQVAPA) at the C-terminal end. This tag can be recognized by the anti-Rho 1D4 antibody, which is useful for protein isolation and tracking their trafficking profile.

Purified DNA plasmids from the RT-PCR were first linearized by *EcoR*1 and *Sal*1 restriction enzyme digestion (New England BioLabs Inc., UK), using 3-5 μg DNA per reaction. The digest were then set up in suitable buffers following the manufacturer's protocol, with an addition of bovine serum albumin (BSA; 0.1 mg/mL; New England BioLabs Inc., UK) to stabilize the enzymes during digestion. Reactions were incubated for at least 12 hours in water bath at 37°C.

After the amplicons were digested with *Eco*RI and *Sal*I restriction enzymes, they were subcloned into mammalian expression vector pMT4 by cohesive-ended ligation. Concentration of the vector and the DNA inserts were determined visually from agarose gels by comparing their intensities with those of the marker DNA fragments. Each ligation reaction was set up to contain 100 ng of vector, 3-fold molar excess of the insert, 1 μL of ligase 10× buffer (Promega, UK), 1 unit of T4 DNA ligase (Promega, UK), and nuclease-free water to make up a final volume of 10 μL. The tubes were incubated at 4°C overnight before transforming into bacteria.

Competent *E.coli* DH5α cells (Life Technologies, UK) were used for transformation and cloning of the plasmids. Following the manufacturer's instructions, 5 μL of each DNA ligation product was mixed with 50 μL of DH5α cells on ice and left incubating for 30 min. The cells were then heat shocked in 42°C water bath for 20 s and placed on ice for 2 min. 950 μL of pre-warmed SOC medium was added to each transformation and the tubes were incubated in shaker at 37°C for 1 h

at 225 rpm. Afterwards, the cells were spun down by centrifugation for 2 min at 800 rpm. 800 μL of the medium was then removed and the cell pellet was resuspended in the remaining ~200 μL medium. Approximately 200 μL of each transformation was spread on pre-warmed selective LB-agar plates containing 100 μg/mL ampicillin. The plates were incubated at 37°C for 16 hours.

To select the correct clone for plasmid preparation, six colonies per gene were picked for sequence screening. Each clone was inoculated in 2 ml LB containing 100 μg/mL ampicillin and grown at 37°C overnight with vigorous shaking. Plasmids were prepared from the cultures with GenElute Plasmid MiniPreps Kit (Sigma-Aldrich, UK) following the manufacturer's protocol and eluted in 50 μL elution solution. Sequencing of the plasmids was performed by DNA Sequencing Lab (Source BioScience LifeScience, Oxford) with pMT4 forward and reverse primers.

After the DNA sequences were analysed, the correct plasmids from MiniPreps were retransformed into 20 μL DH5*α* cells and plated as described in the previous section. Following incubation at 37°C overnight, a single colony was picked from each plate to set up starter cultures of 2 mL LB medium each containing 100 μg/mL ampicillin. The starter cultures were incubated at 37°C with shaking for 6 h. For large-scale plasmid preparation, 500 μL of each starter culture was inoculated in 250 mL LB medium with ampicillin (100 μg/mL), and the suspensions were grown at 37°C for 12-16 h with shaking. Plasmid DNAs were purified using HiSpeed Maxi kit (Qiagen, UK) following manufacturer's instructions and dissolved in 1 mL TE buffer.

2.5.2 Cell line models: maintenance and transfection

Mouse neuroblastoma (Neuro-2A) (ECACC) and Human Embryonic Kidney 293T (HEK-293T) cells were used as model systems for functional characterisation of the tmt opsins. The Neuro-2A cells were grown in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich, UK) supplemented with 10% foetal bovine serum (FBS; Life Technologies, UK), 2 mM L-glutamine (Sigma-Aldrich, UK) and 1% (v/v) penicillin/streptomycin (Sigma-Aldrich, UK). The HEK-293T cells were grown in DMEM containing 4.5 g/L D-glucose and L-glutamine, supplemented with 10% FBS and 1% (v/v) penicillin/streptomycin (Sigma-Aldrich, UK). Both cell lines were kept in incubator at 37°C, with 5% CO_2 and 100% humidity. Media of stock flasks were replaced every 2-3 days, and the cells were routinely passaged before reaching confluence.

All the tmt-pMT4 opsin constructs for transfection were prepared with a plasmid Maxiprep kit (Qiagen, UK). Transfection of Neuro-2A cells was performed using GeneJuice transfection reagent (Novagen) following the manufacturer's protocol. Briefly, Neuro-2A cells were seeded at a density of 2×10^5 cells on cover slips in each 35 mm tissue culture dish. 24 h after seeding, cells were washed with PBS and incubated in RPMI media (Sigma-Aldrich, UK) containing 10% FBS, 2 mM L-glutamine, 2 μg of DNA and 6 μl of GeneJuice reagent for at least 6 h. Cells were then fed normal cell culture media and grown for 24 h before immunocytochemistry (ICC) and functional experiments were performed. Transfection for HEK-293T cells was also performed with GeneJuice, but with 210 μg DNA in 630 μl of the transfection reagent per plate, using twelve 144 mm tissue culture dishes per

experiment (Davies et al., 2012). After 48 h, the transfected cells were harvested, washed four times with PBS and kept in -80°C until needed for spectrophotometric analysis.

2.5.3 *In vitro* spectral analysis of zebrafish tmt opsins

2.5.3.1 Regeneration and purification of tmt opsin proteins

Tmt opsin proteins were expressed in HEK-293T cells as previously described in section 2.6. Plasmid DNA of RH1 1D4 epitope-tagged *tmt2s*, *tmt2l*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, or *tmt24* was transfected into HEK-293T cells using the GeneJuice (see section 2.5.2). After 48 h of cultivation, transfected cells were harvested in 1x PBS (pH 7.0), centrifuged and stored at -80°C until analysis. To regenerate the pigments, cells were resuspended in 1x PBS (pH 7.0) and incubated with 20 μM of either 11-*cis* or 40 μM all-*trans* retinal in the dark at 4°C for 1.5 h. The membrane-bound pigments were then solubilised with buffer containing 1% (w/v) n-Dodecyl-β-D-Maltoside ULTROL® Grade (DDM; Calbiochem, UK) and 0.2 mg/mL protease inhibitor phenylmethylsulfonyl fluoride (PMSF; Sigma-Aldrich, UK) and purified using immunoaffinity chromatography. The purification step was achieved by applying the pigments to a CNBr-activated Sepharose column that was coupled with anti-RH1 1D4 antibody. The column was washed with 1% DDM/PBS buffer and the regenerated pigments were eluted using the same buffer containing a competitive peptide (peptide I) corresponding to C-terminus of bovine rhodopsin (Molday and MacKenzie, 1983, Davies et al., 2007b). The eluent was concentrated by Vivaspin® 2 Centrifugal Concentrator (GE Healthcare, UK) according to the manufacturer's

protocol. Finally, the reconstituted opsin proteins were kept on ice and subjected to analysis by UV-Vis spectrophotometry (as described below).

2.5.3.2 UV-visible spectroscopy

Absorption spectra (200-700 nm) of the regenerated photopigments were recorded in the dark using a Shimadzu UV-2550 UV-Vis spectrophotometer. All spectra were recorded with a bandwidth slit of 1 nm, with either medium or slow scan speed at room temperature. Subsequently, the samples were either photobleached by broad spectrum white light for 1 h or denatured by acidification using 26 mM HCl. The corresponding bleached or acid-denatured spectra were recorded immediately after treatment. Spectrophotometric recordings for each opsin sample were repeated three times. For calculation of the peak absorbance value (λ_{max}) for each expressed opsin protein, beached or acid-denatured spectra were subtracted from the dark absorbance spectra to generate difference spectra. The resultant difference visual spectra were overlaid with a modified visual pigment template (Govardovskii et al., 2000, Parry et al., 2004), and best-fit spectral curves were obtained using an optimisation add-in function called 'Solver' in Excel to vary the λ_{max}. Since absorbance spectra are distorted by the underlying absorbance and scatter of the protein, it is more accurate to use difference spectra as estimations for the λ_{max} values.

2.5.3.3 Investigation of tmt10 photoproduct pH-dependency

To elucidate the mechanisms of tmt opsin activation (using tmt10 pigments as a representative), pH-dependency of its photoproducts was investigated. The

pigments of tmt10 were extracted from HEK293T cells, reconstituted and purified as described before in buffers at testing pH of 5.5 and 9.5. The regenerated photopigments were analysed by UV-Vis spectrophotometry (section 2.5.3.2). λ_{max} of the photoproducts at dark state were deduced from the difference spectrum before and after acid denaturation treatment (see section 2.5.3.2 for detail). By comparing λ_{max} values of different photoproducts, changes in protonation state of the opsin were interpreted.

2.5.3.4 Construction of mutant tmt10 for spectral tuning study

The full-length coding sequence of tmt10 in pMT4 was used as template for the generation of mutant tmt10 constructs by site-directed mutagenesis. Forward and reverse primers carrying one or two mutant nucleotides were designed to induce a single amino acid mutations Tyr113Glu, Glu181Gln, Cys87Ser (nomenclature describes original amino acid, position, and mutation) in *EcoR*I- or *Sal*I-fragments of the tmt10 coding region. Sequences of oligonucleotide primers are listed in table 2.4. First round PCR was set up to amplify the mutant *EcoR*I- and *Sal*I-fragments of tmt10. Each reaction contained 33 μL sterile water, 1× KOD DNA polymerase buffer, 1 mM $MgSO_4$, 0.2 mM dNTP mix, 1 U KOD DNA polymerase, 0.3 μM each forward and reverse primer, and 100 ng tmt10 plasmid DNA template. The PCR was conducted in the following thermal cycles: an initial polymerase activation step at 95°C for 10 min, then 40 cycles of 95°C for 15 s, 55°C for 30 s, and 70°C for 1 min, followed by cooling at room temperature. PCR products of the expected size were excised and purified as described in section 2.4.4.

The amplified fragments were then spliced together (Davies et al., 2007a) by a second round PCR to reconstruct the complete 1.1 kb tmt10 coding region that would encode one or two amino acid substitution(s): Tyr113Glu, Glu181Gln, Cys87Ser, and Tyr113Glu/Glu181Gln. To generate a single continuous tmt10 sequence, only the two outer forward (PE_CHR10F) and reverse (PE_CHR10R) primers were used. The second PCRs were set up to contain: 30 or 31 μL sterile water, 1× KOD DNA polymerase buffer, 1 mM $MgSO_4$, 0.2 mM dNTP mix, 1 U KOD DNA polymerase, 0.3 μM each forward and reverse primer, and equimolar concentration of the first-round PCR products in specific combination (see Fig. 2.2 A-D). Identical thermal cycling conditions of the first round PCR reaction were used for the second PCR, with distinct products excised and purified in the same manner. The purified final PCR products were digested by *EcoR*I and *Sal*I restriction enzymes, and subcloned into the pMT4 expression vector for transformation. Cloning, preparation of plasmids and sequencing were carried out using standard molecular techniques (section 2.5.1). The *tmt10* mutants generated were expressed in HEK293T cells by transfection with GeneJuice, purified, and characterized using UV-Vis spectrometry.

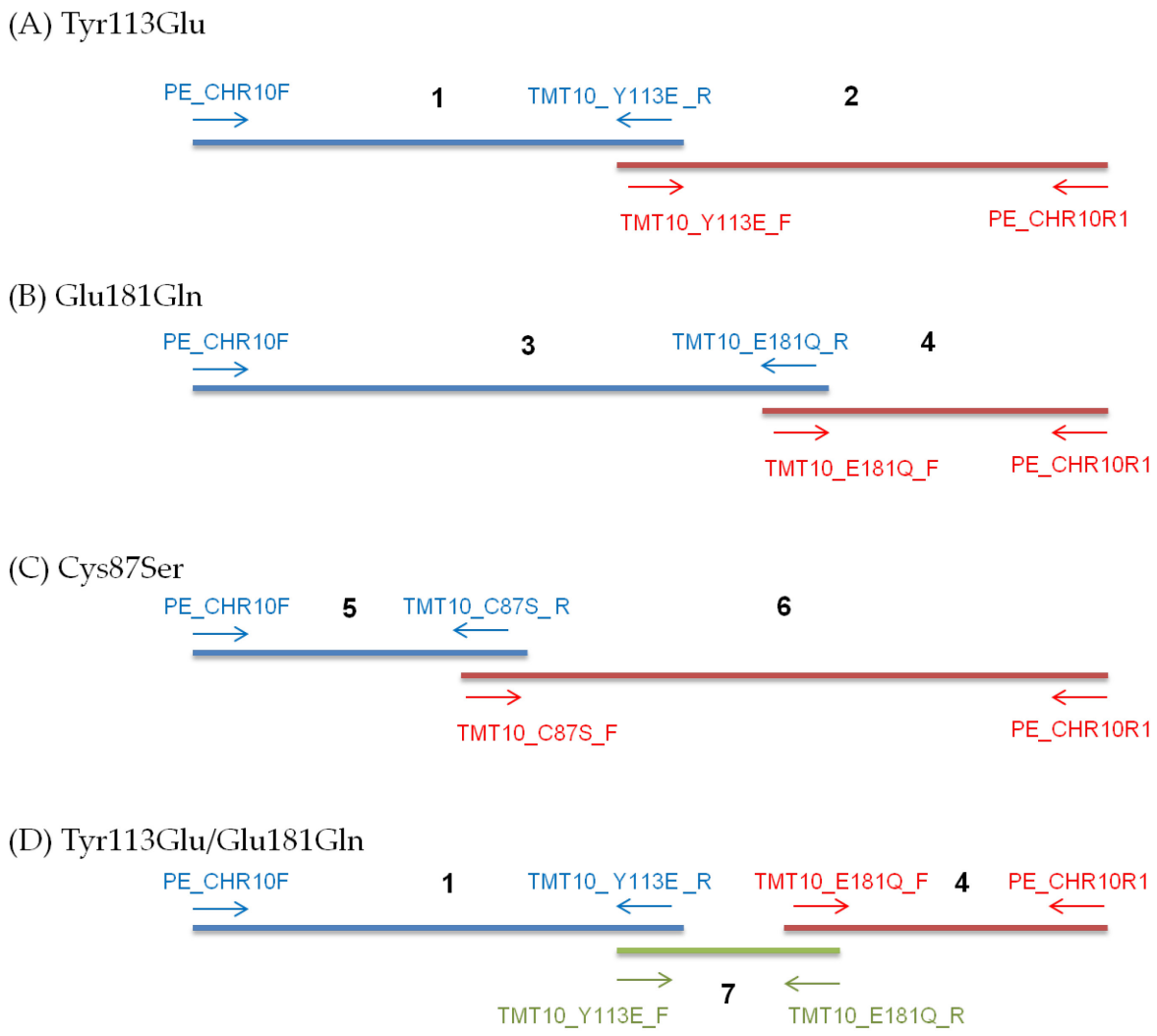


Figure 2.2 Schematic diagrams illustrating the construction of tmt10 mutants carrying site-directed mutations using the primers in table 5.2: **(A)** Tyr113Glu **(B)** Glu181Gln, **(C)** Cys87Ser, and **(D)** Tyr113Glu/Glu181Gln from *Eco*RI- or *Sal*I-fragments.

Table 2.4 *Oligonucleotides used for the generation of zebrafish tmt10 mutants*

Mutations	Fragment	Primers	Oligonucleotide sequences*
Tyr113Glu	1	PE_CHR10F	5′-GCGCGAATTCCACCATGGTCACTGTCCCCTTCCTCC-3′
		TMT10_ Y113E _R	5′- GCGTTGGCAAAACCCTCCCCACCTGCATCCG -3′
	2	TMT10_Y113E_F	5′- CGGATGCAGGTGGGAGGGTTTTGCCAACGC-3′
		PE_CHR10R1	5′-CGGCGTCGACGCTTTGGGGGTGTCCATCAGGCT-3′
Glu181Gln	3	PE_CHR10F	5′-GCGCGAATTCCACCATGGTCACTGTCCCCTTCCTCC-3′
		TMT10_E181Q_R	5′- GTGGTGCCGGGACCCTGAGGACCATAGCTG-3′
	4	TMT10_E181Q_F	5′- CAGCTATGGTCCTCAGGGTCCCGGCACCAC-3′
		PE_CHR10R1	5′-CGGCGTCGACGCTTTGGGGGTGTCCATCAGGCT-3′
Cys87Ser	5	PE_CHR10F	5′-GCGCGAATTCCACCATGGTCACTGTCCCCTTCCTCC-3′
		TMT10_C87S_ R	5′- GGTGTACCGAACAGACTAACCAGCATATCG-3′
	6	TMT10_C87S_F	5′- CGATATGCTGGTTAGTCTGTTCGGTACACC-3′
		PE_CHR10R1	5′-CGGCGTCGACGCTTTGGGGGTGTCCATCAGGCT-3′
Tyr113Glu + Glu181Gln	1	PE_CHR10F	5′-GCGCGAATTCCACCATGGTCACTGTCCCCTTCCTCC-3′
		TMT10_ Y113E _R	5′- GCGTTGGCAAAACCCTCCCCACCTGCATCCG -3′
	4	TMT10_E181Q_F	5′- CAGCTATGGTCCTCAGGGTCCCGGCACCAC-3′
		PE_CHR10R1	5′-CGGCGTCGACGCTTTGGGGGTGTCCATCAGGCT-3′
	7	TMT10_Y113E_F	5′- CGGATGCAGGTGGGAGGGTTTTGCCAACGC-3′
		TMT10_E181Q_R	5′- GTGGTGCCGGGACCCTGAGGACCATAGCTG-3′

* *Eco*RI site = GAATTC (red), *Sal*I site = GTCGAC (blue),

NB: Changes made in the nucleotides for single amino acid substitution (yellow)

2.5.4 Functional assays for identification of tmt opsin signalling pathways

2.5.4.1 Immunocytochemistry (ICC)

Immunofluorescence labelling of tmt opsins tagged by Rho1D4 epitope was performed on transfected Neuro-2A cells. All antibodies were diluted in PBS (Life Technologies, UK)/0.1 Triton X-100 (Life Technologies, UK) supplemented with 2% serum from the same species as the secondary antibody. Briefly, Neuro-2A cells were seeded onto glass coverslips (15 mm diameter; Scientific Laboratory Supplies, UK) in six-well cell culture plates and transfected as described above. 24 h after transfection, cells were fixed with 4% PFA (VWR, UK) for 2 h at room temperature and washed three times with PBS. The cells were then blocked for 1 h in 1 × PBS/0.1% Triton X-100 containing 10% normal donkey serum (Sigma-Aldrich, UK) at room temperature. Subsequent to removal of the blocking solution, the cells were washed with PBS and a primary mouse anti-ID4 monoclonal antibody (Molday and MacKenzie, 1983) (diluted 1:2000 in 1 × PBS with 2% donkey serum and 0.1% Triton X-100) was incubated for 1 h at room temperature. This was followed by another wash of the cells with PBS prior to the incubation with an Alexa 488 donkey anti-mouse secondary antibody (Life Technologies, UK) (diluted 1:500 in 1 × PBS with 2% donkey serum and 0.1% Triton X-100) for 1 h at room temperature. Coverslips were mounted onto slides using ProLong Gold anti-fade reagent with DAPI (Life Technologies, UK) and left for 24 h at room temperature in the dark. Fluorescence images were taken using an Olympus IX71 inverted microscope, excitation 488 nm with emission wavelength 505-530 nm for green fluorescence.

2.5.4.2 Time-lapse calcium imaging

Intracellular Ca^{2+} levels of Neuro-2A cells were monitored in real-time using a cell permeant calcium sensitive dye Rhodamine-2 acetoxymethyl (Rhod-2 AM; Life Technologies, UK). Neuro-2A cells were seeded onto 35 mm Petri dishes and transfected with plasmid as described above. Cells were then treated with 20 μM retinoic acid (RA) in normal DMEM (with 10% FBS, 2 mM L-glutamine and 1% (v/v) penicillin/streptomycin) for 48 h to induce differentiation. Differentiated cells were incubated in media containing 20 μM of either 9-*cis* or all-*trans* retinal for 1.5 h prior to experimentation. Cells were loaded with the dye by incubation with 5 μM Rhod-2 AM, 0.1% Dimethyl sulfoxide (DMSO; Sigma-Aldrich, UK), 0.02% Pluronic F-127 (Life Technologies, UK) and 2.5 mM Probenecid (Life Technologies, UK) for 45 min at 37°C, washed briefly and incubated for a further 20 min in fresh media to remove excess dye. Calcium imaging was performed in Ringer's solution (140 mM NaCl, 4 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM glucose, and 10 mM HEPES; pH 7.3-7.4, 24°C). Rhod-2 AM dye loading and all wash steps were performed under dim red light conditions and cells were incubated in total darkness prior to the onset of any experiment. All cultures were used within 1 h of dye loading. Calcium imaging was performed using an Olympus IX71 inverted microscope fitted with a Xenon arc light source with a slit monochromator (Cairn Optoscan). Images were collected via a high sensitivity CCD camera (Cascade 512B, Photometrics), acquired and analysed using MetaFluor imaging software (Molecular Devices, UK). Visualisation and focusing of the cells prior to imaging was performed briefly under dim red light by placing a 610 nm band pass filter (Chroma Technology, Germany) in the white light path. For time

lapse experiments, Rhod-2 AM fluorescence images were initially recorded every 2 s using a 100 ms exposure of 545 nm light (bandwidth 10 nm, intensity 7.9×10^{13} photons.cm^{-2}.s^{-1}) to excite the Rhod-2 AM dye, with emitted fluorescence collected via a 610 nm band pass filter (Chroma Technology, Germany). All light pulses used for experimental stimulation were performed using a monochromatic light source with the irradiance measured using an optical power meter (Macam Photometrics, UK) and converted to photon flux. Light pulse stimulation was typically in the form of 15 × 1 s exposures of either 420 nm or 480 nm light (bandwidth 20 nm, intensity of 1.1×10^{15} and 1.3×10^{15} photons·cm^{-2}·s^{-1}), interspersed between images 30 and 45 (i.e. 60-90 s). All light stimulus protocols were pre-programmed into and controlled via the MetaFluor imaging software. Analysis of changes in Rhod-2 AM fluorescence was performed by manually defining regions of interest around each cell using MetaFluor Analyst software (Molecular Devices, UK), with data points for individual cells exported to Excel for a detailed analysis.

2.5.4.3 Whole-cell electrophysiological studies of tmt-transfected Neuro-2A cells

Neuro-2A cells were seeded onto glass coverslips (13 mm diameter; Scientific Laboratory Supplies, UK), transfected with plasmids containing a particular tmt coding sequence, differentiated with retinoic acid and loaded with either 9-*cis* or all-*trans* retinal as described previously. Cells were perfused in Ringer's solution (140 mM NaCl, 4 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 5 mM glucose, and 10 mM HEPES; pH 7.3-7.4, 24°C) during recording. Successfully transfected cells were identified by red fluorescence (540 nm excitation filter) of a co-transfected pSIREN-RetroQ-DsRed-

Express empty vector using a Zeiss Axioskop FS2 microscope. Whole-cell patch-clamp recordings were made with pipettes containing intracellular solution (140 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 10 mM HEPES, and 10 mM Egtazic acid (EGTA; Sigma-Aldrich, UK). Osmolarity of the intracellular solution was adjusted to 285 ± 5 mOsmol/L and its pH to 7.3-7.4 with KOH. Open pipette resistance was 2-5 MΩ and access resistance during recordings was less than 20 MΩ. Currents were recorded (Axopatch 200B, Axon Instruments, USA) in neurons voltage-clamped at holding potentials of −50 mV and data were collected using Clampex 10 software (Molecular Devices, UK). The records were filtered at 1 kHz and sampled at 20 kHz. Light stimuli were generated using a Cairn Optoscan Xenon arc source comprising a slit monochromator. All stimuli were 10 s in duration with a 20 nm half-bandwidth. Irradiance was measured using an optical power meter (Macam Photometrics, UK) and converted to photon flux. The magnitude of the responses was defined by the peak sustained current measured and analysed by Clampfit 10 software (Molecular Devices, UK).

2.5.4.4 Luciferase-based bioluminescent assay for screening G_s, G_i or G_q activities

To test which G protein signalling pathway may be activated by tmt opsins in a light-dependent manner, a GloSensor cAMP assay (Promega, UK; for G_i and G_s signal detection) and an aequorin luminescence calcium assay (for G_q signal detection) were performed. Both assays used a vector backbone to express luminescent biosensors encoded with a pcDNA5/FRT/TO vector, which is derived from the FLPN IN system (Life Technologies, UK) and allows expression from a

tetracycline inducible promoter. Both final plasmids of pcDNA5/FRT/TO/Glo22F (for G_i/G_s assays) and pcDNA5/FRT/TO/mt_aq (for G_q assays) were constructed by Dr. H. Bailes in Manchester.

HEK293 cells were plated in 96-well solid white plates at a density of 4 x 105 cells/mL and left overnight. The following day, 200 ng of each zebrafish tmt opsin plasmid was co-transfected with 200 ng of either pcDNA5/FRT/TO/Glo22F or pcDNA5/FRT/TO/mt_aq plasmid into the cells using Lipofectamine 2000 (Life Technologies, UK) in serum-free DMEM for 6 h. Cells post transfection were incubated at 37°C for 16 h in a CO_2 independent medium Leibovitz's L-15 (Life Technologies, UK) without phenol red, in the presence of 10% FBS, 250 ng/mL tetracycline solution (Life Technologies, UK) and 10 μM 9-*cis* retinal (Sigma-Aldrich, UK). All treatments after transfection were performed in dim red light.

For G_s assays, L-15 medium containing 0.1 M Beetle Luciferin, potassium salt (Promega, UK) in 10 mM Hepes buffer was added 16 h after transfection, resulting in 5 μM 9-*cis* retinal in the medium during recording. Cells were equilibrated for 2 h at room temperature and raw luminescence units (RLU) from each well were measured for 1 s every minute with a top-read 3 mm lens in a Fluostar Optima plate reader (BMG Labtech, USA). Subsequent to equilibration, cells were subjected to testing by 5 × 1 s flashes from a PS33 Plus photic stimulator (Grass Technologies, USA) with diffuser removed, followed by 15 min recovery periods where RLU was recorded again. The cells were then exposed to six consecutive light stimulations (i.e. six times of the 5 × 1s flashes, once per min) to test for sustained responses. The emission spectrum of the light source used is shown in Figure 2.4 below.

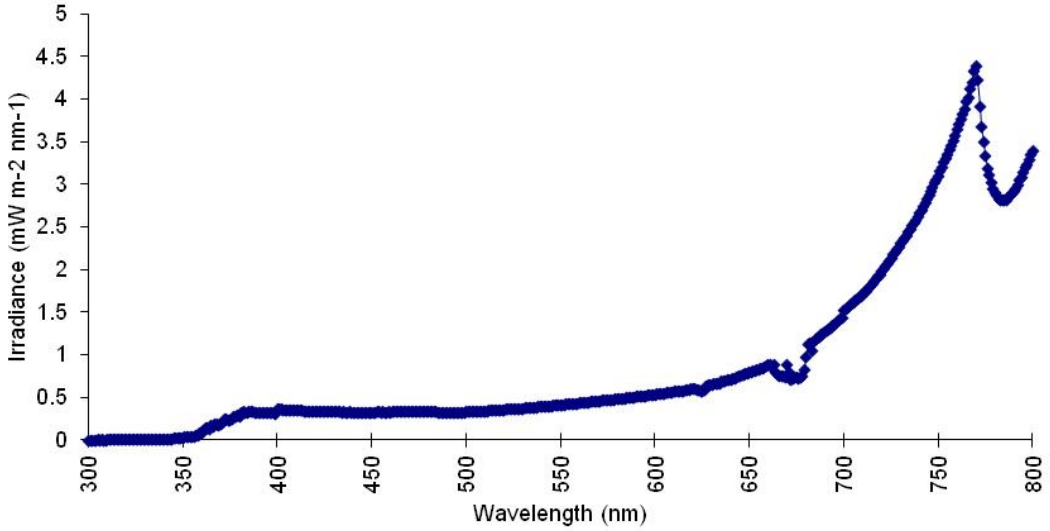


Figure 2.4 Emission spectrum of the light source used for the bioluminescence light response assays. The lamp was set to flash at 20 Hz during measurements, and the spectroradiometer integrates the light over time. Hence, there would be a slight underestimation of the total irradiance of the light source compared to the actual irradiance. This spectra data was collected by Dr. H. Bailes (Manchester, 2012).

For G_i biosensor validation, cells were also treated with 0.1 M Luciferin following transfection. After 2 h equilibration at RT, RLU were recorded from each well at 1 s resolution in the same manner as for the G_s assay. 2 μM forskolin was added to all the wells and RLU were immediately recorded again. 2 min after the addition of forskolin, cells were stimulated by 5 × 1s flashes and then allowed to recover for 15 min. Subsequently, cells were exposed to six consecutive light stimulations for measurement of sustained responses.

For G_q assays, transfected cells were incubated in L-15 medium containing 10 μM coelenterazine h (VWR International, UK) and equilibrated at RT for 2 h. RLU were recorded from individual wells one at a time with 0.5 s resolution. After 15 s of

equilibration RLU were measured, cells were stimulated with 5 × 1 s flashes. Following a recovery period of 35 s, 100 μM of digitonin (Sigma-Aldrich, UK) was added into the well to confirm viability of the cells by releasing their cytosolic calcium store.

The luminescence recording data were analysed using OPTIMA software (BMG Labtech, USA) and Microsoft Office Excel. RLUs recorded from each well were normalised to an average value of the repeats for the 5 min prior to light treatment. The mean of the normalised data were used for generating the light response graphs.

2.5.5 Isolation and characterisation of *Astyanax mexicanus* tmt opsins

2.5.5.1 Astyanax mexicanus (A. mexicanus) cDNA preparation

A. mexicanus cell lines were derived from the embryos of sighted surface form and cave-dwelling forms (Pachón and Steinhardt populations) in Professor Whitmore's laboratory at UCL. Total RNAs were prepared from these cell lines by Dr. Kathy Tamai (UCL). Single-stranded cDNAs were synthesised from the total RNAs using standard methods (section 2.3.2) and stored at -20°C until use.

2.5.5.2 PCR amplification, RACE and cloning of Mexican tetra tmt opsins

Mexican tetra *tmt* opsins partial sequences were initially isolated by degenerate PCR from the cDNA of three different *A. mexicanus* populations (surface, Pachón and Steinhardt). The degenerate primers used were designed by Dr. Wayne Davies to detect sequences specifically from different classes of *tmt* opsins across all teleosts species (see table 2.5). Two successive rounds of PCR were performed to

enhance the sensitivity and specificity in isolating tmt opsin sequences. In all cases, negative controls with no cDNA were included. A first round PCR experiment was consisted of 1× NH_4 buffer, 2 mM $MgSO_4$, 10 μM dNTP mix, 1 μl BIOTAQ DNA polymerase, 40 ng of template cDNA, and 0.2 μM of each F1 forward and R1 reverse primer in RNase-free water. The PCR was performed under the following conditions: an initial denaturation step at 94°C for 5 mins, then 40 cycles of 94°C for 30 s, 45°C for 1 min and 72°C for 1.5 min, ending with an extension step at 72°C for 10 mins. The resulting first round PCR products together with negative controls were diluted 10-fold with sterile water. 1 μL of the diluted samples were then used as templates for a second round of heminested PCR with primer combinations of F1 and R1, F2 and R2, and F1 and R2. Thermal cycling conditions for the second PCR was similar to the first one, but using 50°C for the annealing step.

5′ and 3′ ends of the Mexican tetra *tmt* opsin sequences were obtained by 5′ and 3′ RACE PCR using FirstChoice®RLM-RACE kit (Life Technologies, UK) following the manufacturer's instructions. Primers used were designed based on the partial sequences acquired from degenerate PCR, details of which are listed in table 2.6 and 2.7. 5′ RACE was performed as described before (refer to section 2.4.4.5). 3′ RACE did not require the processing of RNA prior to cDNA synthesis. For 3′ RACE, first strand cDNAs were reverse transcribed directly from total RNAs using the supplied 3′ RACE Adapter. The cDNAs were then subjected to PCR using 3′ RACE outer primer and *tmt*-specific forward primers (F1). 1 μL aliquots of the first round PCR products were then used as template in a second round heminested PCR with 3′ RACE inner reverse primer (R2) and *tmt*-specific forward primer (F2). PCR

conditions for the first and second round PCR were similar to those of described above, except that the annealing temperature was set to 55°C for 1 min.

Full-length tmt6 had been isolated from the surface and Pachón populations of *A. mexicanus* by Andrew Beale (UCL). To confirm their sequences and presence in the Steinhardt population, standard RT-PCR was performed using the following primers: forward 5′-GCGCGAATTCCACCATGTTCTATGATCTGGAGATC-3′ and reverse 5′-CGGCGTCGACGCRGGAGTGATCTGAGGGGGGTGGGTCCGGTC-3′ (method as described in section 2.5.1). These primers were designed to produce full-length sequences that could be easily subcloned into mammalian expression vectors for protein expression experimentation. Note that the reverse primer contained an 'R' symbol at one of the nucleotide positions, which means that it can be an A or a G at this site. The design of this primer takes into account of codon degeneracy, whereby each amino acid can be encoded for by more than one codon, with different base at any of the three positions. Such incorporation of sequence ambiguity allows the generation of a set of 'mixed' primers to accommodate for all the possible codons for the amino acid located at that site. The set of degenerate base symbols that has been used universally is shown in below in Table 2.8.

For both degenerate and RACE PCRs, the first and second round amplified products were visualised by ethidium bromide-stained agarose gel electrophoresis, excised and purified as previously described. The amplicons were directly cloned into a pGEM®-T easy cloning vector (Promega, UK) by TA cloning using T4 DNA ligase. Ligated plasmids were transformed into competent *E.coli* DH5*α* cells, which were grown in LB-agar plates containing 100 μg/mL ampicillin, 0.2 mM IPTG and 40

μg/mL X-gal. Blue/white screening of the colonies allowed for identification of recombinant clones carrying inserts (white colonies). Twelve of these white colonies were picked for inoculation of starter cultures, which were then used to grow larger cultures for scale-up (as described in section 2.4.5). Subsequent isolation of the plasmids was performed by using GenElute Plasmid MiniPreps Kit (Sigma-Aldrich, UK) according to the manufacturer's instructions and confirmed by test digestion with *EcoR*1 restriction enzyme. The selected positive clones were sent with M13 forward primer (5′-CACGACGTTGTAAAACGAC-3′) (Promega, UK) and amplicon-specific reverse primers to Source BioScience LifeScience (Oxford, UK) for sequencing. Multiple alignments of the *Astyanax* partial and full-length tmt opsin sequences against those of zebrafish were performed using ClustalW Multiple Alignment software program. The identities of the opsin fragments obtained were confirmed by blast search against the zebrafish nucleotide genome database using blastn program (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Table 2.5 *Degenerate primers used to isolate partial cDNA of A. mexicanus tmt opsins*

Gene Specificity	Primers	Oligonucleotide sequences
Class 1	TEL_TMT_1_F1	5′-TAGTGCTCGTGYTKTTYTGYAARTTYAAGA-3′
	TEL_TMT_1_F2	5′-GCCSGTKAACATGCTKYTDCTKAACATCAG-3′
	TEL_TMT_1_R1	5′-RTADATGAGDGGRTTGATGACBGTGCTGCT-3′
	TEL_TMT_1_R2	5′-SMGGAAACACYTGTARAACTGYTTGTTCAT-3′
Class 2	TEL_TMT_2_F1	5′-GCTGACGGGACATCACGGVTGYVRRTGGTA-3′
	TEL_TMT_2_F2	5′-TGGCTCCTGGCTCTACKCNCTSBYMTGGAC-3′
	TEL_TMT_2_R1	5′-CAGCGCCAYSAYBCCRTANGGCATCCAGCA-3′
	TEL_TMT_2_R2	5′-CACATAGATGACDGGGTTVASVRCVGTGCT-3′
Class 3	TEL_TMT_3_F1	5′-GGTGGTGGCGGTGTKYYTBGGMTTBATCGG-3′
	TEL_TMT_3_F2	5′-TGGTCWGCGTDCTSGGVACBCCSTTCAGCT-3′
	TEL_TMT_3_R1	5′-CAGGGCCATCAYNCCRTANGGSARCCAGCA-3′
	TEL_TMT_3_R2	5′-CCDRAAACAYCTGYARAACTGTTTRTTCAT-3′

These degenerate primers were designed by Wayne I. L. Davies

Table 2.6 *Primer sequences used in 3′-RACE of Astyanax tmt opsins*

Nested primer name	Primer sequence (5′ to 3′)	Description
3′-RACE Adapter	GCGAGCACAGAATTAATACGACT CACTATAGGT12VN	Anneal to 5′-end of decapped mRNA
3′-RACE Outer primer	GCGAGCACAGAATTAATACGACT	Nested 5′-RACE outer reverse primers
3′-RACE Inner primer	CGCGGATCCGAATTAATACGACT CACTATAGG	Nested 5′-RACE inner reverse primer
3′-RACE_PC_tmt9_F1	TGGACGCCGTACAGCGTGGTGGC CATAATG	3′-RACE tmt9 outer forward primer
3′-RACE_PC_tmt9_F2	CGTTTGGTCGGCCAGGAATCATC ACACCTG	3′-RACE tmt9 inner forward primer
3′-RACE_PA/SH_tmt14_F1	TAAAGGGGTGACTAAGATCAATCT GCTGTC	3′-RACE tmt14 outer forward primer
3′-RACE_PA/SH_tmt14_F2	GCGGGAGAACCACATTCTACTGA TGGTGCT	3′-RACE tmt14 inner forward primer
3′-RACE_PC_tmt24_F1	TCGGGCATCCTGGCCTGGTCACT CCAGAAG	3′-RACE tmt24 outer forward primer
3′-RACE_PC_tmt24_F2	TTCCCTCATTACTCGCCAAATCCA GCACAG	3′-RACE tmt24 inner forward primer
DEGEN_3′-RACE_tmt10/14_F1	GCCACTCCTCCTCATGRTSTWCT GCTATGG	Degenerate 3′-RACE tmt10/14 outer forward primer
DEGEN_3′-RACE_tmt10/14_F2	GAGGAGGGAGAACCACRTKCTDS TGATGGT	Degenerate 3′-RACE tmt10/14 inner forward primer
DEGEN_3′-RACE_TMT2/24_F1	ATCCTGGACTGGTGACDSCRGMA GCCAGCA	Degenerate 3′-RACE tmt2/24 outer forward primer
DEGEN_3′-RACE_TMT2/24_F2	ATCCTGGACTGGTGACDSCRGMA GCCAGCA	Degenerate 3′-RACE tmt2/24 outer forward primer

Table 2.7 *Primer sequences used in 5′-RACE of Astyanax tmt opsins*

Nested primer name	Primer sequence (5′ to 3′)	Description
5′-RACE Adapter	GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA	Anneal to 5′-end of decapped mRNA
5′-RACE Outer primer	GCTGATGGCGATGAATGAACACTG	Nested 5′-RACE outer forward primers
5′-RACE Inner primer	CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG	Nested 5′-RACE inner forward primer
5′-RACE_PC_tmt9_R1	AATCCGTACCACATGCAGCCCTGCCTCCCA	5′-RACE tmt9 outer reverse primer
5′-RACE_PC_tmt9_R2	GTGGATGCTGGACGCGAAGCTGAGCGTCGT	5′-RACE tmt9 inner reverse primer
5′-RACE_PA/SH_tmt14_R1	TGCTGGGGGATCTCTGGTGCCACTGCACAG	5′-RACE tmt14 outer reverse primer
5′-RACE_PA/SH_tmt14_R2	GGTGGTTCCTGGACCTTCAGGCCCATAGCT	5′-RACE tmt14 inner reverse primer
5′-RACE_PC_tmt24_R1	CGGCCAGGGAGATCAGAGACACGATACCCA	5′-RACE tmt24 outer reverse primer
5′-RACE_PC_tmt24_R2	AGTTGACGAAGCCGTACCACACGCAGCCAG	5′-RACE tmt24 inner reverse primer
DEGEN_5′-RACE_tmt10/14_R1	TTAATGATAAAAAGGATYTTGCCATAGCAG	Degenerate 5′-RACE tmt10/14 outer reverse primer
DEGEN_5′-RACE_tmt10/14_R2	TCCGGGACCTTCAGGMCCATAGYTGCTCCA	Degenerate 5′-RACE tmt10/14 inner reverse primer
DEGEN_5′-RACE_TMT2/24_R1	GATTCCCATGATCACYTTYYKGTAGTTGGT	Degenerate 5′-RACE tmt2/24 outer reverse primer
DEGEN_5′-RACE_TMT2/24_R2	GAACGAGTTGACGAARCCGTACCACACGCA	Degenerate 5′-RACE tmt2/24 outer reverse primer

Table 2.8 *Standardized nomenclature of nucleotide degenerate codes (IUPAC)*

Degenerate base symbols	Represented bases
W	A, T
S	C, G
M	A, C
K	G, T
R	A, G
Y	C, T
B	C, G, T
D	A, G, T
H	A, C, T
V	A, G, C
N	A, G, C, T

CHAPTER 3

Discovery of Zebrafish tmt Opsin Gene Family

CHAPTER 3: Discovery of Zebrafish tmt Opsin Gene Family

3.1 Introduction

Light detection is one of the most important mechanisms in determining the immediate environment surrounding an organism. It may help to promote the survival of the organism by facilitating its search for energy sources and mate selection. Although light is essential for vision, it also constitutes a powerful modulator for a broad range of non-image-forming functions that includes circadian photoentrainment, pupil constriction, locomotion and masking behaviours (Berson, 2003, Brainard and Hanifin, 2005, Rollag et al., 2003, Gamlin et al., 2007, Foster and Bellingham, 2002). The circadian system, in turn, controls many other biological activities that encompasses but is not limited to hormone synthesis, feeding behaviour, and sleep regulation (Rollag et al., 2003, Hatori and Panda, 2010, Dijk and Archer, 2009).

The hypothesis that most organs, and in fact tissues of zebrafish, contain irradiance detectors was supported by the identification of a novel non-visual photopigment in teleost fish, the teleost multiple tissue (tmt) opsin (Moutsaki et al., 2003). Initial studies by Moutsaki et al. showed that tmt opsin is expressed within and external to the CNS, as well as in an embryonic cell-line (PAC2) of zebrafish. The PAC2 cell-line has been demonstrated to exhibit circadian rhythmicity in its *clock* gene expression when the cells are placed in a light-dark cycle (Whitmore et al., 2000). Photoentrainable clock gene expression was also observed by Whitmore and colleagues in majority of the organs in zebrafish. Furthermore, comparison of the deduced amino acid sequence of tmt opsin with rod and cone opsins revealed

conserved features, including a retinal chromophore binding site (Lys296) that is suggestive of photoreceptive function (Moutsaki et al., 2003). Collectively, the expression pattern and structural characteristics of tmt opsin indicated that this opsin might act as photopigment mediating peripheral circadian oscillations. The objectives of this chapter are therefore to: (1) establish the full complement of tmt opsins that are present in the zebrafish genome, (2) analyse their structural properties and (3) examine their phylogenetic relationships with other known opsins.

3.2 Results

3.2.1 Cloning of zebrafish tmt opsins reveals new subclasses of tmt opsins

The published zebrafish *tmt* opsin nucleotide sequence (AF349947) is a 765 bp fragment that encodes a single partial putative opsin protein of 255 aa. This sequence encompasses part of the amino- or N-terminal domain and seven predicted transmembrane regions of the tmt opsin, containing all of the essential features required for photopigment structure and function (Moutsaki et al., 2003). These features include Lys296 as a binding site for retinal chromophore, Cys110 and Cys187 forming a disulphide bond to stabilise the opsin, and an E(D)RY motif which also stabilises the opsin molecule when it is inactive. By performing 3′-rapid amplification of cDNA ends (3′-RACE) and reverse transcription polymerase chain reaction (RT-PCR) amplifying both ends of the known *tmt* opsin, a full-length version of the coding region with complete amino (N)- and carboxyl (C)-termini was obtained (Wayne I. L. Davies, unpublished data). Comparison between the two sequences showed that they possessed the same first three exons, but differed in the

fourth exon. Closer analysis indicated that the published *tmt* opsin is a "short" isoform splice variant which lacks the fourth exon. Instead, it has an extension of exon 3 into intron 3 to the next available stop codon, giving it a much shorter C-terminal tail than the long isoform. Variation in the length and sequence of C-terminal tail is commonly observed with other opsins like opn4 and VA opsin, some of which have been shown to have isoform-specific functions (Kojima et al., 2000, Minamoto and Shimizu, 2002, Pires et al., 2009, Davies et al., 2012, Hughes et al., 2012). A BLAST search by nucleotide blast (blastn) using the nucleotide sequences of these two opsin genes as bait against the zebrafish genome confirmed that they are part of the same gene, located on chromosome 2 and did not arise from a gene duplication or gene conversion event. Basic Local Alignment Search Tool (BLAST) is a bioinformatic tool used by researchers to perform sequence similarity searches by comparing a nucleotide, deduced amino acid, or protein sequence in query against a database of sequences.

Bioinformatic screening of the *tmt* opsin nucleotide sequences revealed six *tmt*-like orthologues of this gene in the zebrafish genome, and their full-length expressions were verified using RT-PCR (see Section 2.3.3). The complete cDNA sequences of each *tmt* opsins are included in Appendix A.1. A blastn search of these nucleotide sequences indicated that they are discretely located on individual chromosomes (Chr 2, Chr 6, Chr 9, Chr 10, Chr 14 and Chr 24), hence they are tentatively called *tmt2l (long)*, *tmt2s (short)*, *tmt6*, *tmt9*, *tmt10*, *tmt14* and *tmt24*. Previously performed, yet unpublished genomic analysis showed that apart from the short isoform of *tmt2*, all the *tmt* opsins isolated consist of 4 exons (Fig. 3.1). The sizes

of the exons/introns and nucleotides around the splice sites are shown in Table 3.1. Comparison of exon sizes between the six sequences showed that exon 2 and 3 are of similar length in different *tmt* opsins. However, the sizes of exon 1 and 4 vary significantly between the genes, due to the variable length of the N- and C-termini, a common feature of all vertebrate opsin classes (Wayne I. L. Davies, unpublished data). All the exon/intron splice junctions obey the gt/ag rule, which is required for spliceosomes to recognise and process the transcribed pre-mRNA into mature mRNA (Mount, 1982). Analysis of nucleotide sequences at the donor splice sites (5′ end of the intron), the acceptor splice sites (3′ end of the intron), and in particularly the branch sites of each splice junction showed high degree of identities to their consensus sequences, with averages of 81%, 82%, and 92% similarities respectively (see Table 3.2). The sizes of open reading frames predicted from these sequences are as follow: *tmt2l* – 1164 bp, *tmt2s* – 966 bp, *tmt6* – 1113 bp, *tmt9* – 1146 bp, *tmt10* – 1074 bp, *tmt14* – 1140 bp and *tmt24* – 1194 bp.

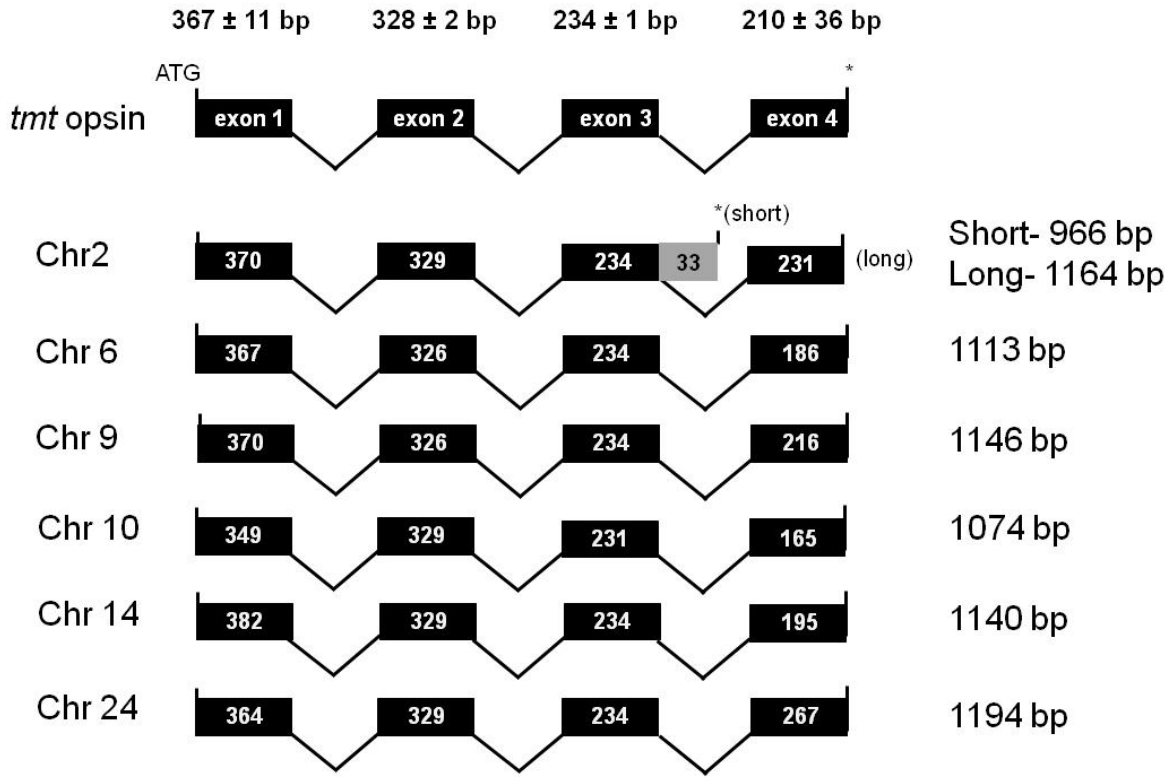


Figure 3.1 The genomic organisation of zebrafish tmt opsins. Exons are shown as black boxes and intron as lines; the sizes of each are marked as base pair (bp) within the boxes. The average length of each exon is indicated with standard deviation above the consensus gene structure on the top row, whilst the sizes of predicted open reading frames are marked at the end of each tmt opsin on the right. The start (ATG) and stop (*) codons of each genes are indicated in the representative *tmt* opsin on the top row. All tmt opsin transcripts encompass 4 exons, with tmt2 generated an additional short splice variant. This splice variant possesses exons 1-3 only, with exon 3 running into intron 3 (grey box) to the next available stop codon located within the intron.

Table 3.1 *Exon and intron organisation of zebrafish tmt-opsin genes*

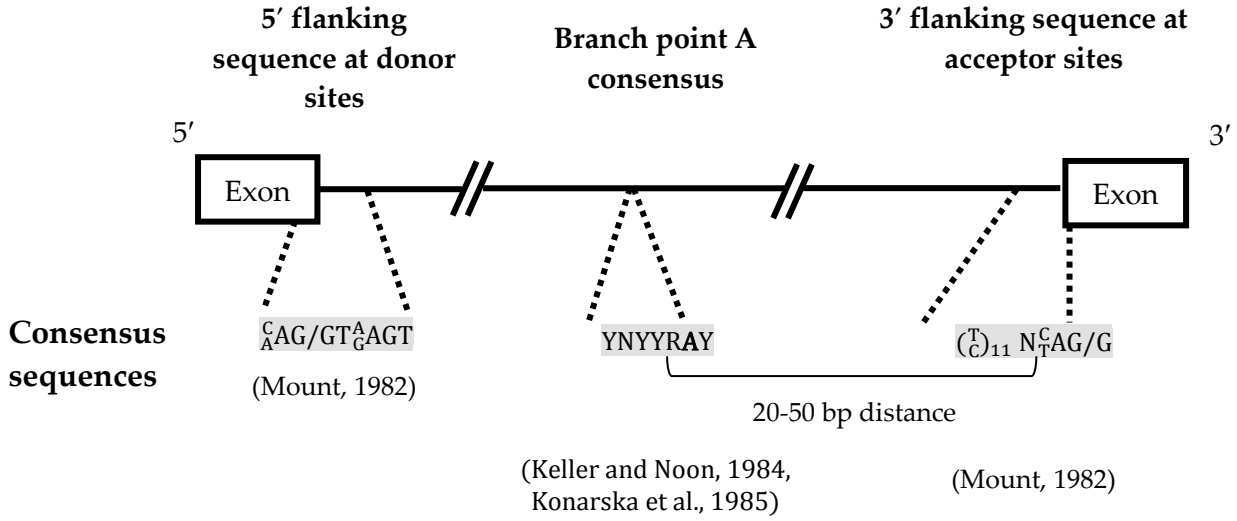
	Gene sequences (5′-3′)*	Exon size (bp)	Intron size (bp)
tmt2	ATGATT……TGTTAGgtgagt…	370	1028
	…ttgcagGCATTG……AAACAGgtgagt…	329	49236
	…atgaagGTCAGC……GGATTTgtgagc…	234	24127
	…gttaagTTCTAC……GGATGA	231	
tmt6	ATGTTT……GCTTTGgtgagt…	367	46902
	…tgacagGGATTG……AAACAGgtaaat…	326	8430
	…ccacagGTAGGC……AAACAGgtgaga…	234	12309
	…gttcagTTTTAC……GGCTAA	186	
tmt9	ATGTTT……GCTTCGgtacgt…	370	51550
	…tcttagGTATTG……AAACAGgtaaac…	326	25071
	…tctcagGTAGGG……AAACAGgcaaga…	234	1115
	…ctgcagTTCTAC……CCATGA	216	
tmt10	ATGGTC……TGTTCGgtaagc…	349	1012
	…caacagGTATCG……CATGGGgtgagt…	329	1232
	…ccacagGTTGCT……AATCAGgtatgc…	231	68
	…tgccagTTTTAC……AAATGA	165	
tmt14	ATGGTC……TTTTTGgtaagt…	382	4033
	…aaacagGAATTG……AAAGGGgtgagt…	329	6082
	…gctcagGTAACA……AACCAGgtgagt…	234	2471
	…tttcagTTCTAC……CCCTGA	195	
tmt24	ATGATT……TCCTGGgtgagt…	364	7964
	…gtacagGCGTTG……ACACAGgtaggt…	329	35545
	…acaaagGTGAGC……AAACAGgtaggt…	234	16858
	…tctcagTTTTGC……GGCTGA	267	

*Introns are in lowercase, exons are in uppercase and boxed

*Key features are highlighted: start codons = yellow; stop codons = green, gt/ag splice sites = gray)

*gt = donor site, ag = acceptor site

Table 3.2 *Comparison of acceptor, donor splice sites and region surrounding branch point A in the pre-mRNAs of tmt opsins to the consensus sequences.*



tmt2	...TAG/GTGAGT... (89%)	...AACCA**A**T... (86%)	...TTAAATTGTTTGCAG/G... (75%)
	...CAG/GTGAGT... (100%)	...TACTG**A**T... (100%)	...TATCTGATGATGAAG/G... (63%)
	...CAG/GTGAGC... (89%)	...CTCTA**A**T... (100%)	...TGTGTGTGTGTTAAG/T... (56%)
tmt6	...TTG/GTGAGT... (89%)	...CTCTG**A**T... (100%)	...TTTTTTTTTTGACAG/G... (94%)
	...CAG/GTAAAT... (89%)	...TCTGA**A**T... (86%)	...TTTGTTTTTCCACAG/G... (94%)
	...CAG/GTGAGA... (89%)	...CACTA**A**T... (100%)	...CAGTGTCTTGTTCAG/T... (69%)
tmt9	...TCG/GTACGT... (67%)	...TTCTG**A**C... (100%)	...TCTTCCCTCTCTTAG/G... (100%)
	...CAG/GTAAAC... (78%)	...TATTG**A**C... (100%)	...GTCTCTCTTTCTCAG/G... (94%)
	...CAG/GCAAGA... (78%)	...TGTCA**A**T... (100%)	...TCTCTTCTCCTGCAG/T... (94%)
tmt10	...TCG/GTAAGC... (67%)	...CATAA**A**T... (86%)	...TTCCAACATCAACAG/G... (75%)
	...GGG/GTGAGT... (78%)	...TTGTT**A**C... (71%)	...CTCCTCATCCCACAG/G... (94%)
	...CAG/GTATGC... (78%)	...TGTCA**A**T... (100%)	...GGCGTTTGCTGCCAG/T... (63%)
tmt14	...TTG/GTAAGT... (78%)	...TGACA**A**T... (86%)	...CTTTACACCAAACAG/G... (75%)
	...GGG/GTGAGT... (78%)	...GCTTA**A**T... (100%)	...TCCATTTTGGCTCAG/G... (81%)
	...CAG/GTGAGT... (100%)	...TGGTG**A**T... (86%)	...CTTCATCTGTTTCAG/T... (88%)
tmt24	...TGG/GTGAGT... (78%)	...TTATA**A**C... (86%)	...CCTGTGTTGGTACAG/G... (75%)
	...CAG/GTAGGT... (89%)	...TCCTC**A**T... (86%)	...CCTCATCACACAAAG/G... (75%)
	... CAG/GTAGGT... (89%)	...TCTCA**A**T... (100%)	...GTCTTTGTGTCTCAG/T... (81%)

*The numbers in the brackets represent the percentage identity agreement with consensus sequence

** Y = any pyrimidine (C or T); R = any purine (A or T); **A** = branch point; $(^{C}_{T})_{11}$ = pyrimidine tract

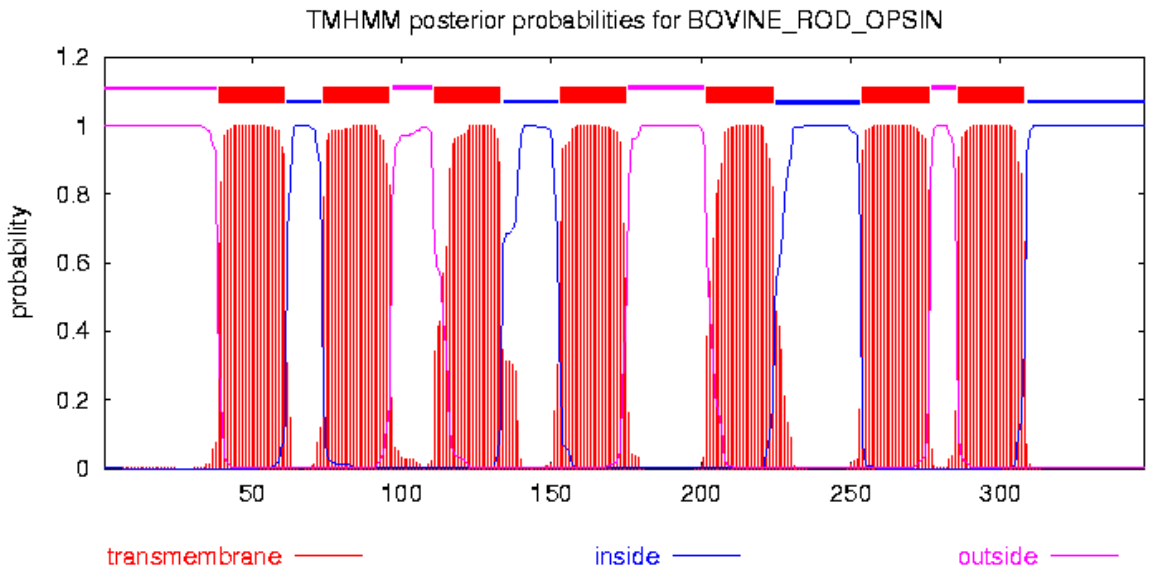
3.2.2 Key structural features of zebrafish tmt opsins

The predicted amino acid sequences of tmt opsins, and their most closely related sister group zebrafish opn3 (sequence determined by Wayne I. L. Davies, unpublished), were analysed to investigate the residues that are important for structural features and opsin function. Similar to tmt opsins, opn3 is an opsin that was found to be expressed in multiple tissues (Halford et al., 2001a). It is important to note that the analyses in this section are based on rod visual pigment, thus some of the features may not relate to non-visual opsins. Nonetheless, by comparing the conserved features of well-characterised visual opsin with those of tmt opsins may provide an insight into their functions.

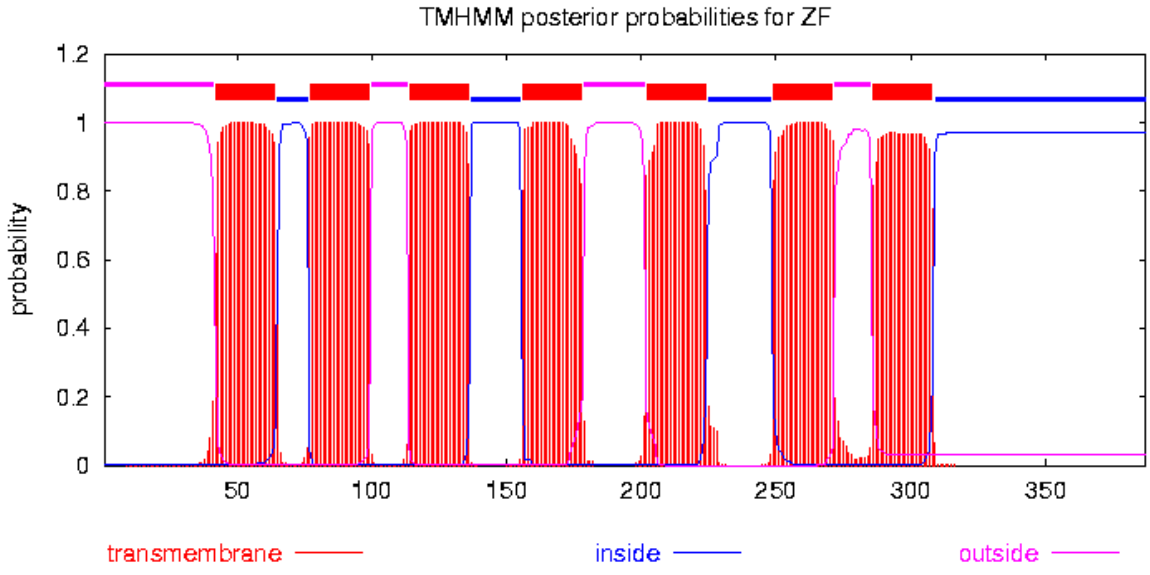
Putative amino acid sequence of each tmt opsin was subjected to an online secondary structure prediction program (TMHMM Server v. 2.0, http://www.cbs.dtu.dk/services/TMHMM-2.0/), which deduced the topology of the proteins based on their hydrophobicity index, charge bias, helix lengths, and topological constrains (Krogh et al., 2001; see Table. 3.3). Bovine rod opsin (RH1; NP001014890) was included in the analysis as a control comparison. Both programs predicted seven transmembrane regions (TMs) for all seven tmt opsin proteins (see Table 3.3), with three cytoplasmic loops (CLs), three extracellular loops (ECs), an extracellular N-terminus and an intracellular carboxyl C-terminus at locations similar to those of the bovine RH1, which were determined by X-ray crystallography (Palczewski et al., 2000). This suggests that tmt opsins share the same membrane topology as the GPCR superfamily, specifically within the family A receptors that includes all other retinylidene proteins (Kristiansen, 2004).

Table 3.3 *Hydrophobicity plots of TM prediction for bovine RH1 and zebrafish tmt opsins*

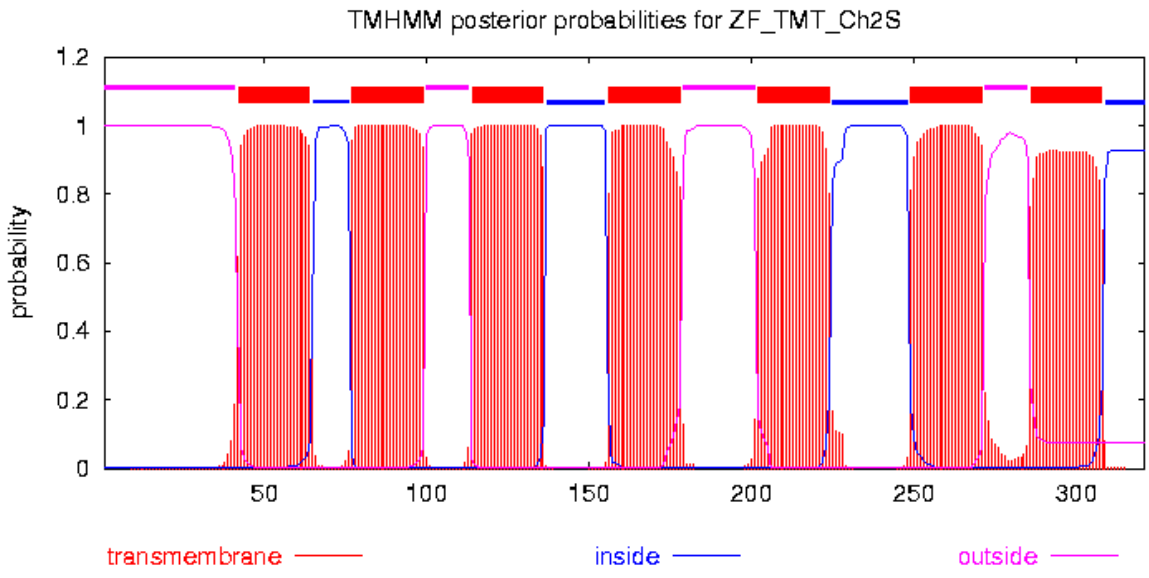
(A) Bovine RH1



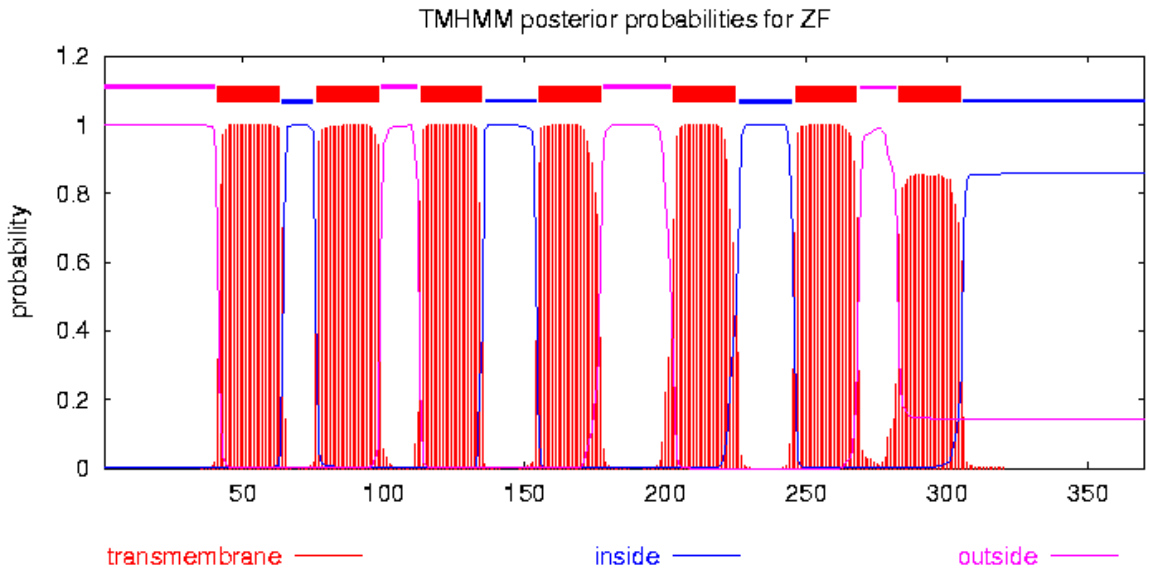
(B) tmt2l



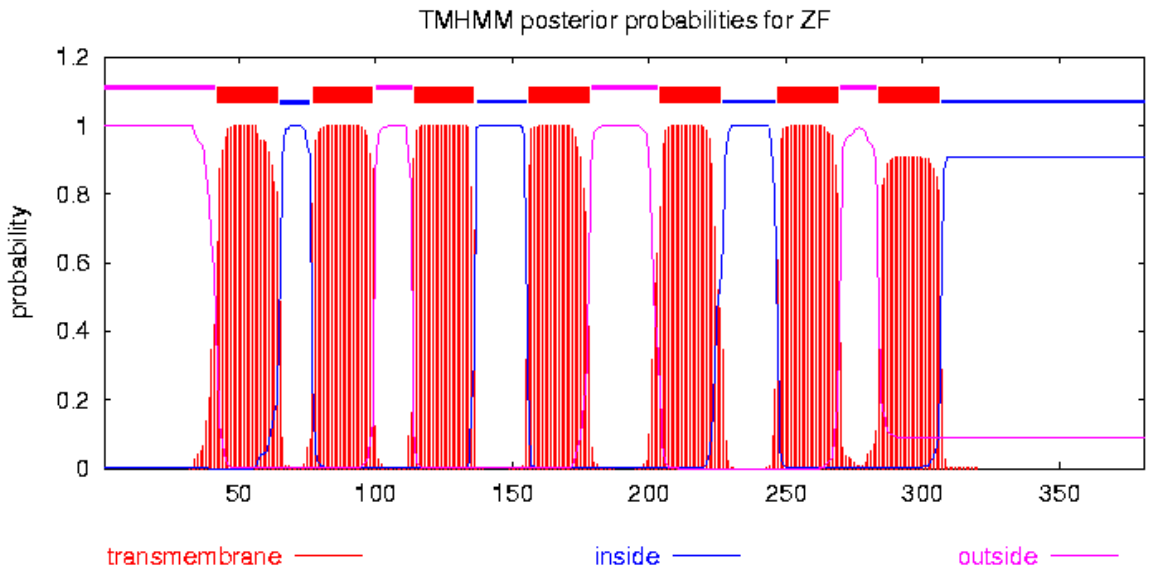
(C) tmt2s



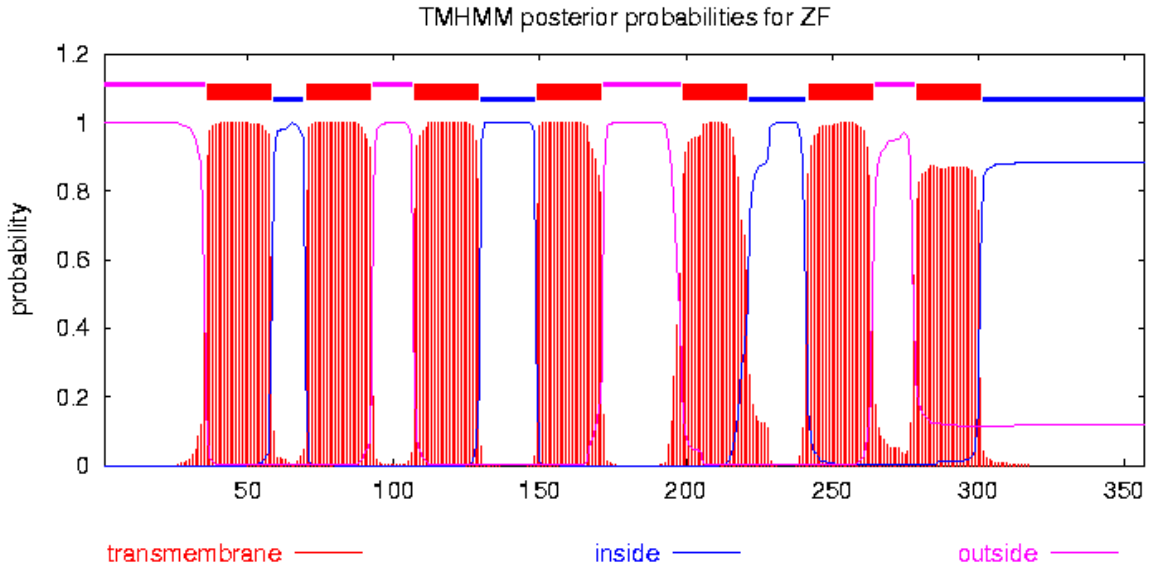
(D) tmt6



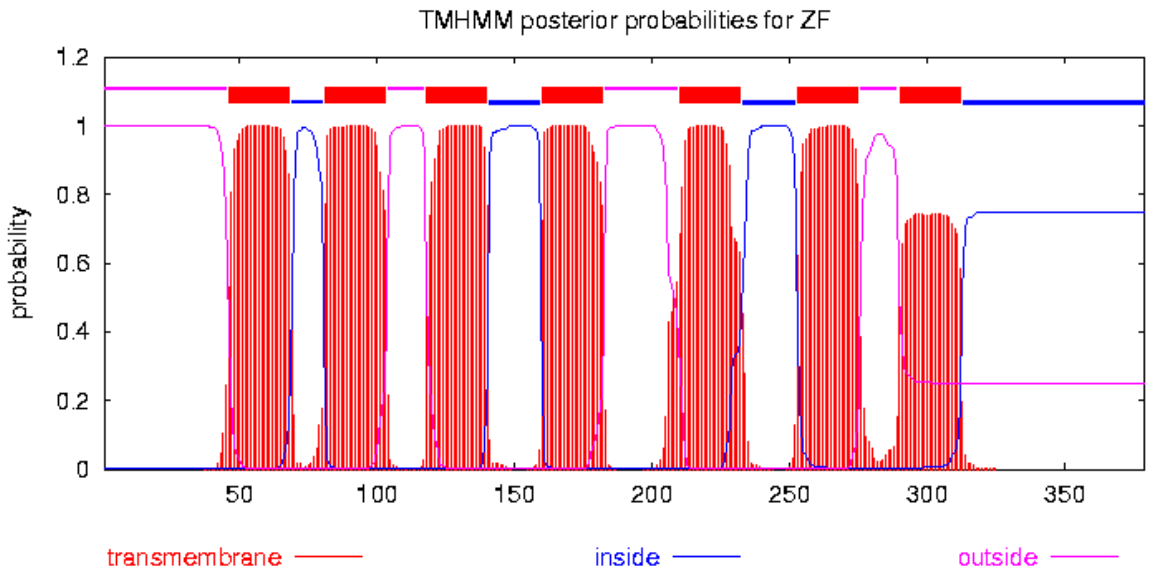
(E) tmt9



(F) tmt10



(G) tmt14



(H) tmt24

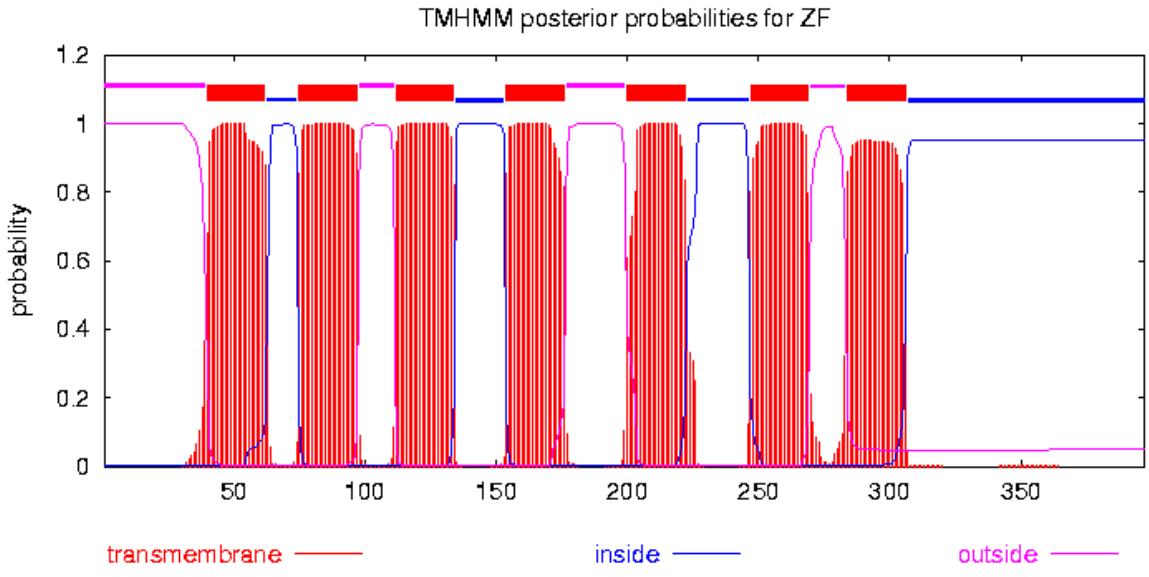


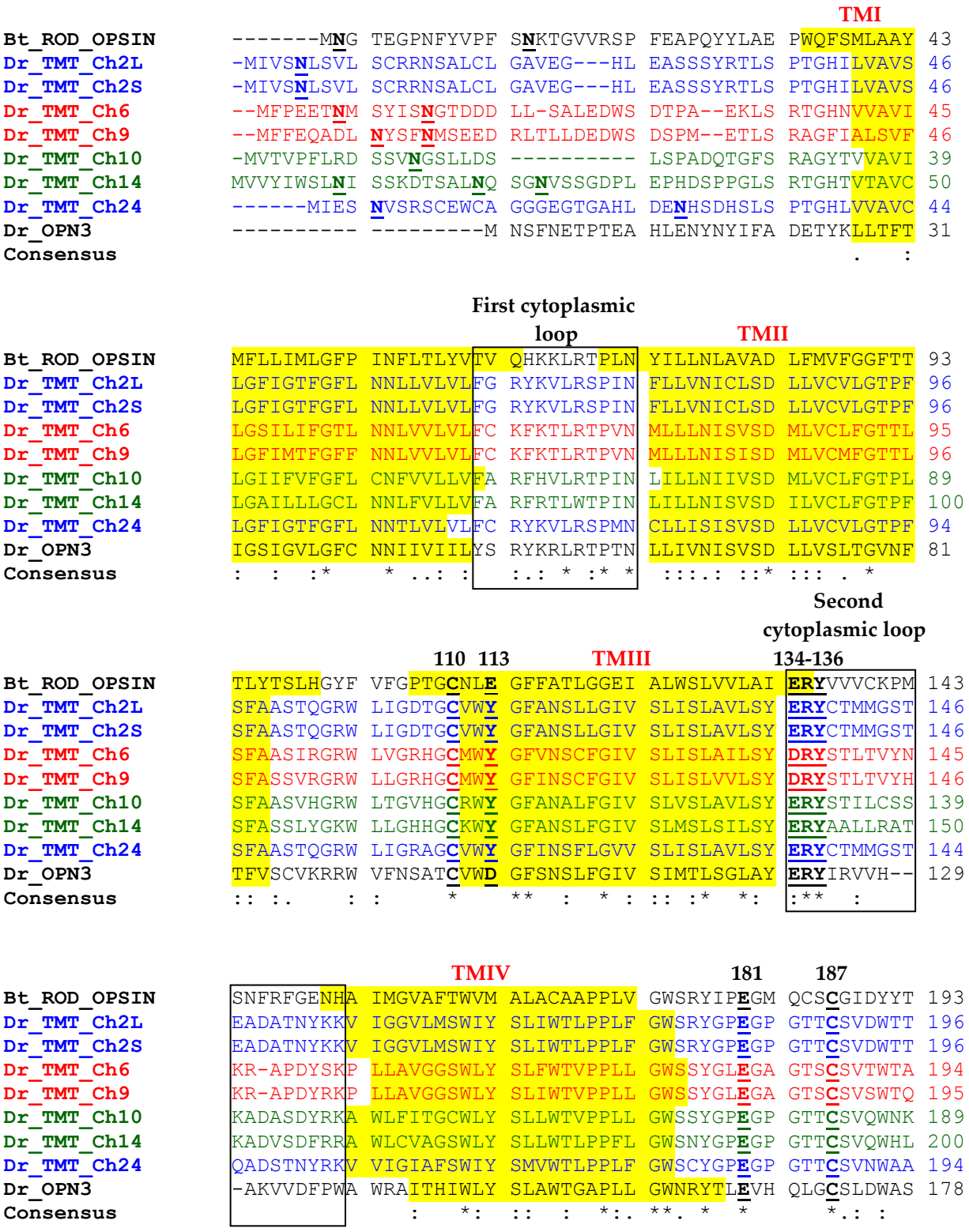
Table 3.3 Hydrophobicity plots of *Danio rerio* tmt opsins. The deduced amino acid sequences of **(A)** bovine RH1, **(B)** tmt2l **(C)** tmt2s, **(D)** tmt6, **(E)** tmt9, **(F)** tmt10, **(G)** tmt14, and **(H)** tmt24 were analysed using an online program TMHMM Server v2.0 (http://www.cbs.dtu.dk/services/TMHMM/) for prediction of hydrophobic transmembrane regions. Probability of the tmt sequences forming transmembrane domains are indicated in red lines, ranging between 0 - 1.0 on the y-axis. The intracellular domains are shown in blue lines, whilst the extracellular domains are shown in pink lines.

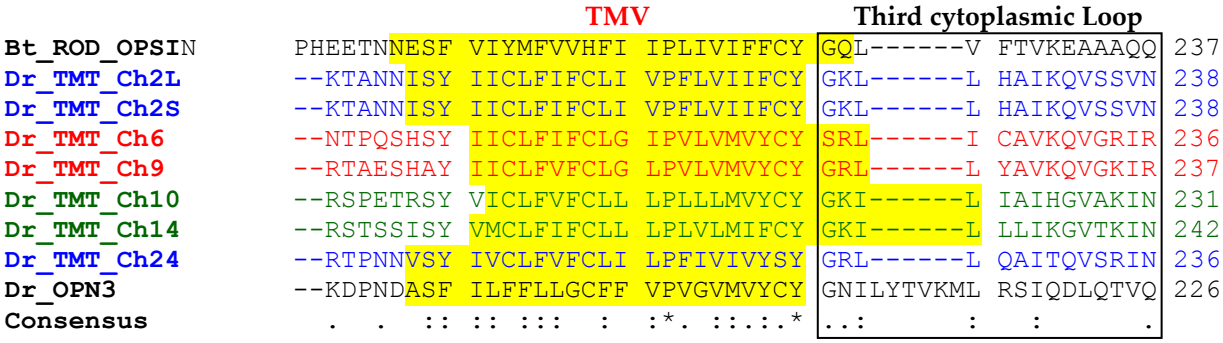
In general, the core regions are relatively conserved across the opsin sequences (with ~65% amino acid identity), but the N- and C-termini differ significantly (with only 22% and 19% amino acid identities respectively). By comparing the amino acid sequences of all seven tmt opsin proteins and opn3 to the well-characterised bovine rod opsin (RH1) protein sequence, it was demonstrated that most of the critical residues known to be required for optimal tertiary structure and function of opsin protein are present (as summarised in Fig. 3.2B). Using the conventional numbering scheme based on the bovine rod opsin polypeptide sequence, these key residues (shaded grey and light blue in Fig. 3.2A) are as follow: (a) a conserved Lys296 in TM7 which serves as a binding site for retinal via a protonated (or unprotonated in UV-sensitive visual pigments (Hunt et al., 2004)) Schiff base linkage (Hargrave et al., 1983); (b) a pair of conserved cysteine residues (Cys110 and Cys187) in EC2 and EC3 that helps to maintain conformational stability (Karnik et al., 1988); (c) a highly conserved ERY [E(D)/R/Y(W/F)] motif at the boundary of TM3 and CL2 (position 134-136), where it has been shown to be essential for G protein activation (Franke et al., 1992); (d) conservation of one or two cysteine residue(s) (Cys322 and Cys323), which are putative palmitoylation sites at the C-terminus in all tmt opsins and opn3, except for tmt10 and tmt2s (Ovchinnikov Yu et al., 1988a) (predicted online using CSS-Palm, http://csspalm.biocuckoo.org/); (e) the NPxxY$(x)_{5,6}$F motif in TM7, which interacts with the ERY motif to mediate structural changes required for activation of G protein (Fritze et al., 2003); (f) the presence of one or two potential glycosylation (aspartate [N]) sites at the N-terminus, but not conserved in location (Hargrave, 1977) (predicted using online software

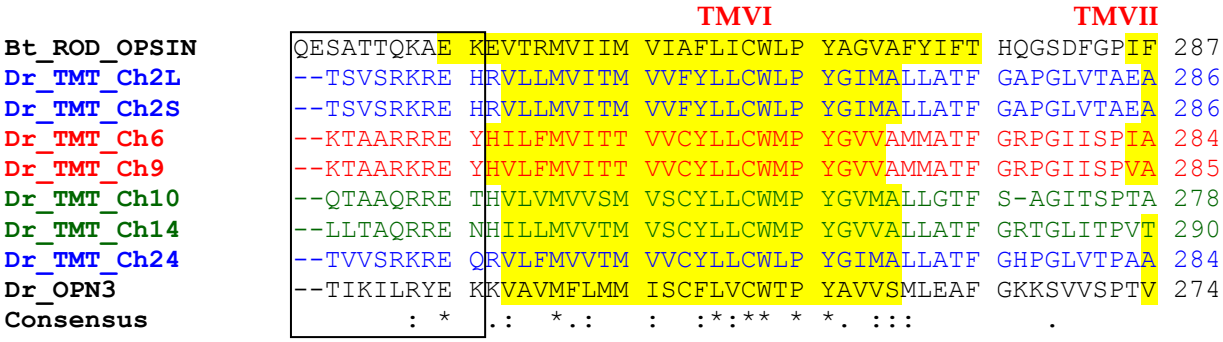
NetNGlyc Server 1.0, http://www.cbs.dtu.dk/services/NetNGlyc/) ; and (g) enrichment of threonine (T) and serine (S) residues at the C-terminus as potential phosphorylation sites (predicted by GPS2.1 Online Service program for Phosphorylation) targeted by rhodopsin kinases for deactivating metarhodopsin II (Palczewski et al., 1993).

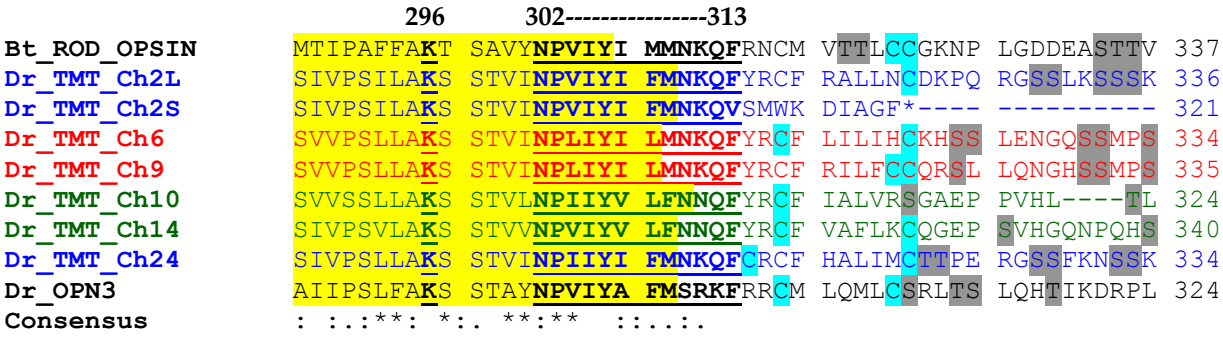
Whilst tmt opsins and opn3 share many features with classical visual pigments, several differences are observed. One notable characteristic is the amino acid located at position 113, which is an uncharged tyrosine in all tmt opsins and a negatively charged aspartate in opn3. In visual pigments, the counterion is typically a negatively charged glutamate (E) at site 113. Since aspartate carries the same charge as glutamate, it may also function as a counterion (Sakmar et al., 1989). A tyrosine substitution at this position is often found in many non-cone/non-rod opsins (e.g. opn4 and opn5) (Gartner and Towner, 1995, Bellingham et al., 1998). It has been suggested that Glu181, which is highly conserved in all but two opsin classes (LWS and VA opsins), may act to counter-balance the positive charge of the Schiff base linkage in the absence of Glu113 (Terakita et al., 2004). Based on this, it seems that tmt opsins may also use Glu181 as a counterion if indeed their Schiff base is protonated. It has been shown that a counterion to the Schiff base linkage is maintained even in pigments with unprotonated Schiff base, such as UVS opsins (Kusnetzow et al., 2004, Hunt et al., 2004). This is because a protonation event of Schiff base linkage is still required for the activation and subsequent hydrolysis of the UVS opsins during photoactivation.

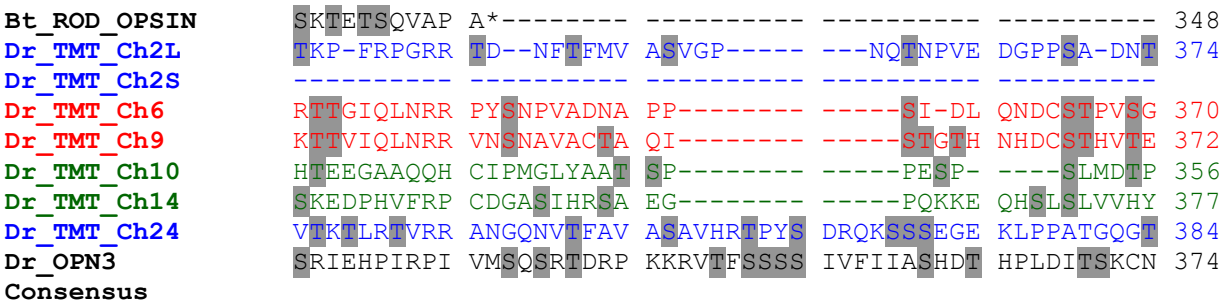
(A)

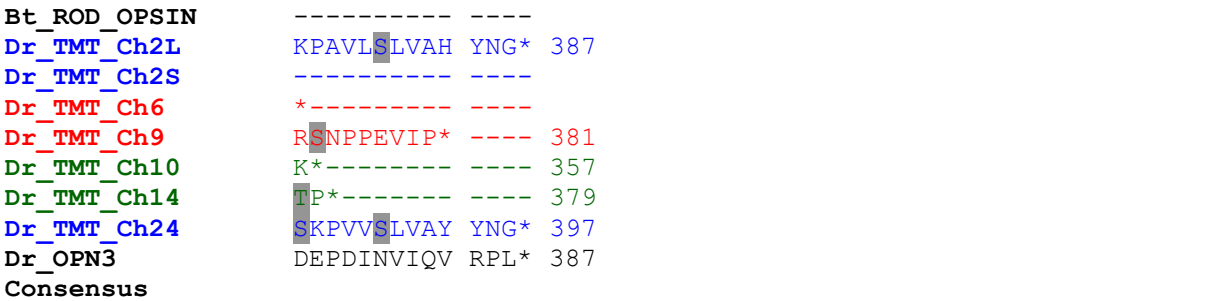

TMI
First cytoplasmic loop
TMII
Second cytoplasmic loop
110 113
TMIII
134-136
TMIV
181 187
Bt_ROD_OPSIN
Dr_TMT_Ch2L
Dr_TMT_Ch2S
Dr_TMT_Ch6
Dr_TMT_Ch9
Dr_TMT_Ch10
Dr_TMT_Ch14
Dr_TMT_Ch24
Dr_OPN3
Consensus


TMV
Third cytoplasmic Loop
Bt_ROD_OPSIN PHEETNNESF VIYMFVVHFI IPLIVIFFCY GQL------V FTVKEAAAQQ 237
Dr_TMT_Ch2L --KTANNISY IICLFIFCLI VPFLVIIFCY GKL------L HAIKQVSSVN 238
Dr_TMT_Ch2S --KTANNISY IICLFIFCLI VPFLVIIFCY GKL------L HAIKQVSSVN 238
Dr_TMT_Ch6 --NTPQSHSY IICLFIFCLG IPVLVMVYCY SRL------I CAVKQVGRIR 236
Dr_TMT_Ch9 --RTAESHAY IICLFVFCLG LPVLVMVYCY GRL------L YAVKQVGKIR 237
Dr_TMT_Ch10 --RSPETRSY VICLFVFCLL LPLLLMVYCY GKI------L IAIHGVAKIN 231
Dr_TMT_Ch14 --RSTSSISY VMCLFIFCLL LPLVLMIFCY GKI------L LLIKGVTKIN 242
Dr_TMT_Ch24 --RTPNNVSY IVCLFVFCLI LPFIVIVYSY GRL------L QAITQVSRIN 236
Dr_OPN3 --KDPNDASF ILFFLLGCFF VPVGVMVYCY GNILYTVKML RSIQDLQTVQ 226
Consensus


TMVI
TMVII
Bt_ROD_OPSIN QESATTQKAE KEVTRMVIIM VIAFLICWLP YAGVAFYIFT HQGSDFGPIF 287
Dr_TMT_Ch2L --TSVSRKRE HRVLLMVITM VVFYLLCWLP YGIMALLATF GAPGLVTAEA 286
Dr_TMT_Ch2S --TSVSRKRE HRVLLMVITM VVFYLLCWLP YGIMALLATF GAPGLVTAEA 286
Dr_TMT_Ch6 --KTAARRRE YHILFMVITT VVCYLLCWMP YGVVAMMATF GRPGIISPIA 284
Dr_TMT_Ch9 --KTAARKRE YHVLFMVITT VVCYLLCWMP YGVVAMMATF GRPGIISPVA 285
Dr_TMT_Ch10 --QTAAQRRE THVLVMVVSM VSCYLLCWMP YGVMALLGTF S-AGITSPTA 278
Dr_TMT_Ch14 --LLTAQRRE NHILLMVVTM VSCYLLCWMP YGVVALLATF GRTGLITPVT 290
Dr_TMT_Ch24 --TVVSRKRE QRVLFMVVTM VVCYLLCWLP YGIMALLATF GHPGLVTPAA 284
Dr_OPN3 --TIKILRYE KKVAVMFLMM ISCFLVCWTP YAVVSMLEAF GKKSVVSPTV 274
Consensus


296
302----------313
Bt_ROD_OPSIN MTIPAFFAKT SAVYNPVIYI MMNKQFRNCM VTTLCCGKNP LGDDEASTTV 337
Dr_TMT_Ch2L SIVPSILAKS STVINPVIYI FMNKQFYRCF RALLNCDKPQ RGSSLKSSSK 336
Dr_TMT_Ch2S SIVPSILAKS STVINPVIYI FMNKQVSMWK DIAGF*---- ---------- 321
Dr_TMT_Ch6 SVVPSLLAKS STVINPLIYI LMNKQFYRCF LILIHCKHSS LENGQSSMPS 334
Dr_TMT_Ch9 SVVPSLLAKS STVINPLIYI LMNKQFYRCF RILFCCQRSL LQNGHSSMPS 335
Dr_TMT_Ch10 SVVSSLLAKS STVLNPIIYV LFNNQFYRCF IALVRSGAEP PVHL----TL 324
Dr_TMT_Ch14 SIVPSVLAKS STVVNPVIYV LFNNQFYRCF VAFLKCQGEP SVHGQNPQHS 340
Dr_TMT_Ch24 SIVPSLLAKS STVINPIIYI FMNKQFCRCF HALIMCTTPE RGSSFKNSSK 334
Dr_OPN3 AIIPSLFAKS STAYNPVIYA FMSRKFRRCM LQMLCSRLTS LQHTIKDRPL 324
Consensus


Bt_ROD_OPSIN SKTETSQVAP A*-------- ---------- ---------- ---------- 348
Dr_TMT_Ch2L TKP-FRPGRR TD--NFTFMV ASVGP----- ---NQTNPVE DGPPSA-DNT 374
Dr_TMT_Ch2S ---------- ---------- ---------- ---------- ----------
Dr_TMT_Ch6 RTTGIQLNRR PYSNPVADNA PP-------- -----SI-DL QNDCSTPVSG 370
Dr_TMT_Ch9 KTTVIQLNRR VNSNAVACTA QI-------- -----STGTH NHDCSTHVTE 372
Dr_TMT_Ch10 HTEEGAAQQH CIPMGLYAAT SP-------- -----PESP- ----SLMDTP 356
Dr_TMT_Ch14 SKEDPHVFRP CDGASIHRSA EG-------- -----PQKKE QHSLSLVVHY 377
Dr_TMT_Ch24 VTKTLRTVRR ANGQNVTFAV ASAVHRTPYS DRQKSSSEGE KLPPATGQGT 384
Dr_OPN3 SRIEHPIRPI VMSQSRTDRP KKRVTFSSSS IVFIIASHDT HPLDITSKCN 374
Consensus


Bt_ROD_OPSIN ---------- ----
Dr_TMT_Ch2L KPAVLSLVAH YNG* 387
Dr_TMT_Ch2S ---------- ----
Dr_TMT_Ch6 *--------- ----
Dr_TMT_Ch9 RSNPPEVIP* ---- 381
Dr_TMT_Ch10 K*-------- ---- 357
Dr_TMT_Ch14 TP*------- ---- 379
Dr_TMT_Ch24 SKPVVSLVAY YNG* 397
Dr_OPN3 DEPDINVIQV RPL* 387
Consensus

(B)

Conserved features	**AA position**	***Bt* Rh1**	***Dr* tmt2l**	***Dr* tmt2s**	***Dr* tmt6**	***Dr* tmt9**	***Dr* tmt10**	***Dr* tmt14**	***Dr* tmt24**	***Dr* opn3**
Putative glycosylation sites*	Varies	2	1	1	2	2	1	3	2	0
Conserved Glu or Tyr	113	E	Y	Y	Y	Y	Y	Y	Y	D
Conserved Cys pair	118, 187	✓	✓	✓	✓	✓	✓	✓	✓	✓
E(D)/R/Y(W/F) motif	134-136	ERY	ERY	ERY	DRY	DRY	ERY	ERY	ERY	ERY
Conserved Glu	181	✓	✓	✓	✓	✓	✓	✓	✓	✓
NPxxY(x)$_{5,6}$F motif	203-313	✓	✓	✓	✓	✓	✓	✓	✓	✓
Retinal binding site	296	K	K	K	K	K	K	K	K	K
Putative palmitoylation sites	322, 323	✓	✓	✓	✓	✓	✓	✓	✓	✓
Potential phosphorylation sites * [threonine, serine]	Varies	9	13	1	12	15	8	8	19	17

* Numbers of potential glycosylation and phosphorylation sites are indicated.

Table 3.4 *Comparison of nucleotide and amino acid sequence identities for the third cytoplasmic loop of zebrafish tmt opsins with that of zebrafish opn3 and bovine RH1.*

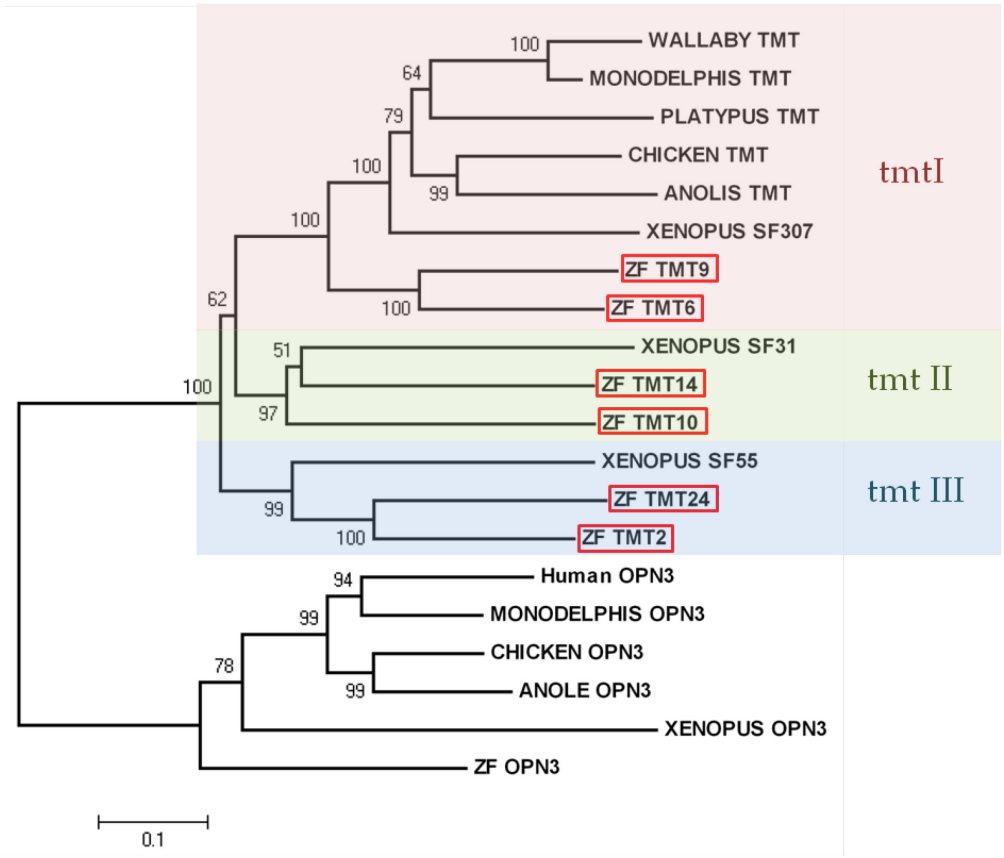
		Nucleotide sequence identity (%)							
	Tmt	***tmt2*** [III]	***tmt6*** [I]	***tmt9*** [I]	***tmt10*** [II]	***tmt14*** [II]	***tmt24*** [III]	**Bovine *RH1***	***opn3***
Amino acid identity (%)	**tmt2** [III]	100	49	61	45	49	68	43	38
	tmt6 [I]	35	100	62	51	43	48	37	26
	tmt9 [I]	48	78	100	52	55	61	45	33
	tmt10 [II]	39	39	48	100	64	43	39	32
	tmt14 [II]	39	30	39	65	100	51	41	36
	tmt24 [III]	70	43	52	39	35	100	36	36
	Bovine *RH1*	20	16	24	16	20	16	100	34
	opn3	24	35	48	39	39	17	13	100

Sequence identities were calculated in BioEdit Sequence Alignment Editor program (Hall, 1999). The upper right numbers (shaded pink) are the percentage identities of nucleotide sequences of the opsins, whilst the lower left numbers are the percentage identities of deduced amino acid sequences of the opsins. The superscript numbers next to subheading of each tmt opsin represent their classes.

3.2.3 Phylogenetic analysis of zebrafish tmt opsins

The phylogenetic relationship between zebrafish tmt opsins and other known opsins found in other vertebrate species is shown in Figure 3.3. The trees were constructed using a neighbour-joining method (Saitou and Nei, 1987) on codon-matched alignment of nucleotide and amino acid sequences (see Section 2.3.6 for detailed methods) (Wayne I. L. Davies, unpublished data). Phylogenetic analyses show that tmt opsins clade into three independent orthologous gene subclasses (tmtI, tmtII and tmtIII). *Xenopus tropicalis* (frog) possess three *tmt* opsin genes, one in each of the subclasses. In teleosts such as zebrafish, two copies of *tmt* opsin in each class are present: *tmt9* and *tmt6* in the tmt1 group, *tmt10* and *tmt14* in the tmt2 group, and *tmt24* and *tmt2* in the tmt3 group. These copies are resulted from a putative whole genome duplication, which occurred early in the teleost fish lineage (Taylor et al., 2001, Meyer and Schartl, 1999, Hoegg et al., 2004). Interestingly, tmtII and tmtIII were lost in the amniotes, and tmtI was subsequently lost in the eutherian mammals. Broader phylogenetic analysis shows that tmt opsins are closely related to opn3, sharing ~87% amino acid identity with one another. However, opn3 forms a separate monophylogenic clade distinct from the tmt opsin family, which indicates that they are two completely different groups of opsins. This is important because historical data suggested they were orthologues, with opn3 being present in mammals whilst tmt opsins in non-mammals (White et al., 2008). However, current data demonstrate that opn3 is also present in zebrafish (*Danio rerio*) (Wayne I.L. Davies, unpublished). Findings from this study on tmt opsins may therefore provide functional clues on the role of its close relative opn3.

(A)


WALLABY TMT
MONODELPHIS TMT
PLATYPUS TMT
CHICKEN TMT
ANOLIS TMT
XENOPUS SF307
ZF TMT9
ZF TMT6
XENOPUS SF31
ZF TMT14
ZF TMT10
XENOPUS SF55
ZF TMT24
ZF TMT2
Human OPN3
MONODELPHIS OPN3
CHICKEN OPN3
ANOLE OPN3
XENOPUS OPN3
ZF OPN3
100
64
79
100
99
100
62
100
51
100
97
99
100
94
99
78
99
0.1
tmtI
tmt II
tmt III

(B)

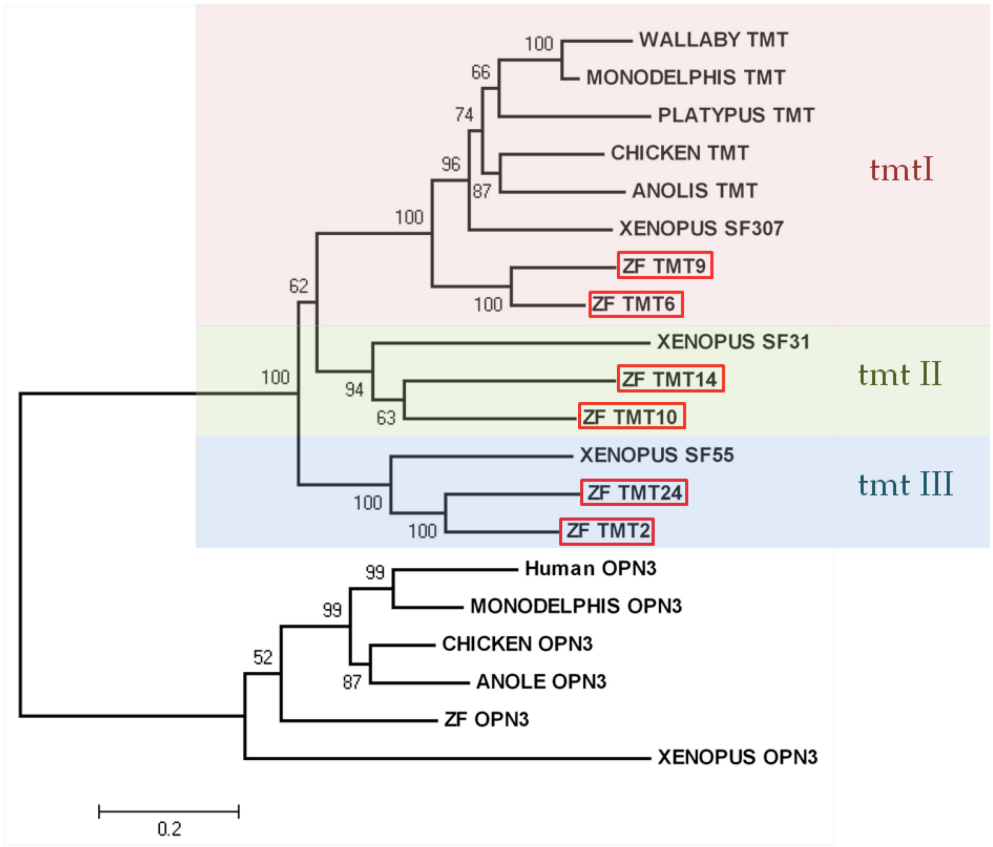

WALLABY TMT
MONODELPHIS TMT
PLATYPUS TMT
CHICKEN TMT
ANOLIS TMT
XENOPUS SF307
ZF TMT9
ZF TMT6
XENOPUS SF31
ZF TMT14
ZF TMT10
XENOPUS SF55
ZF TMT24
ZF TMT2
Human OPN3
MONODELPHIS OPN3
CHICKEN OPN3
ANOLE OPN3
ZF OPN3
XENOPUS OPN3
100
66
74
96
87
100
62
100
100
94
63
100
100
99
99
52
87
0.2
tmtI
tmt II
tmt III

Figure 3.3 Phylogenetic analyses of tmt opsins (class I, shaded pink; class II, shaded green; and class III, shaded blue) from a range of vertebrate species (Wayne I. L. Davies, unpublished data). The six tmt opsins isolated from zebrafish are highlighted in red boxes. The neighbour-joining trees were generated based on codon-matched sequence alignments of **(A)** the nucleotide sequences using Maximum Composite Likelihood (MCL) (Tamura and Nei, 1993),and **(B)** the amino acid sequences using Jones-Taylor-Thornton (JTT) substitution matrix (Jones et al., 1992). The trees were constructed with 1000 bootstrapping replicates. Internal branch supports are expressed as percentages at the node of each branch. The scale bar indicates the respective number of nucleotide and amino acid substitution per site. Refer to Section 2.3.6 for detailed methodology.

Table 3.5 shows the percentage of identity for both nucleotide and amino acid sequences between the six full-length coding sequences of zebrafish *tmt* opsins. These results demonstrate that *tmt* opsins of the same class have higher similarities in their nucleotide and protein sequences compared to those in different classes (e.g. tmt10 and tmt14 of tmtII shares 58% amino acid identity, but only 44-50% with other tmt opsins). Based on the overall low percentage identities amongst zebrafish *tmt* opsins, this supports the phylogenetic analysis that a gene duplication event might have occurred early in evolution of the ray-finned fish lineage, just before the divergence of teleost. Of all the *tmt* opsins, *tmt9* and *tmt6* share the highest identity in both amino acid (76%) and nucleotide (70%) sequences. This could be due to zebrafish-specific duplication for tmt class I. Without many different fish to compare with, it is possible that these subclass duplication events may have occurred not at the start of the teleost radiation, but within the subfamilies. However, it seems unlikely as duplication of a different tmt class (tmtII) has also been found in other teleosts like *fugu rubripes* (Wayne I. L. Davies, unpublished data). One possible explanation for this is that the two *tmt* opsin genes of tmtI could have been under more intense selective pressure compared to those of the other classes to remain conserved. This might be why their ancestral non-duplicated orthologue was maintained in some mammals, whereas the others were lost.

Table 3.5 *Nucleotide and amino acid sequence identities between the six full-length zebrafish tmt opsins.*

		Nucleotide sequence identity (%)					
	Tmt	***tmt2*** **III**	***tmt6*** **I**	***tmt9*** **I**	***tmt10*** **II**	***tmt14*** **II**	***tmt24*** **III**
Amino acid identity (%)	**tmt2 III**	100	46	46	45	48	68
	tmt6 I	51	100	76	52	54	51
	tmt9 I	54	70	100	53	54	50
	tmt10 II	51	50	49	100	58	50
	tmt14 II	55	49	47	58	100	51
	tmt24 III	66	46	45	44	44	100

Sequence identities of both nucleotides and amino acids were calculated in BioEdit Sequence Alignment Editor program (Hall, 1999). The upper right numbers (shaded pink) are the percentage identities of nucleotide sequences of *tmt* opsins, whilst the lower left numbers are the percentage identities of deduced amino acid sequences of tmt opsins. The superscript numbers next to subheading of each tmt opsin represent their classes.

3.3 Discussion

Using bioinformatics and molecular approaches, new members of the *tmt* gene family were identified. In the zebrafish genome, these comprise six different members- *tmt2*, *tmt6*, *tmt9*, *tmt10*, *tmt14* and *tmt24*. At least one of the *tmt* opsins, *tmt2*, exists as both long (4 exons) and short (3 exons) splice variants, differing only at the C-terminal end (Wayne I. L. Davies, unpublished data). This C-terminal region is known to interact with receptor kinases (Schertler, 1998) and has been implicated for protein trafficking (Concepcion et al., 2002). It may, therefore, be postulated that the short isoform may have altered membrane expression and/or different phototransduction activity, which may be confirmed by immunocytochemistry (ICC) and downstream functional assays (see Chapter 6). The primary structures of these six *tmt* opsin genes (Section 3.2.2), their genomic organization (Table 3.1), and phylogenetic relationship with tmt opsins identified from other species (Section 3.2.3) were examined in detailed and discussed below.

In teleost fish, the family of tmt opsins was identified with broad pattern of tissue expression outside of the CNS, strongly supporting the hypothesis that they are the circadian photopigments for peripheral clocks in zebrafish (Moutsaki et al., 2003). The discovery of multiple tmt opsins in zebrafish has extended the number of subclasses within this gene family, all of which are closely related to opn3 (Halford et al., 2001b). Comparison of the amino acid sequences of zebrafish tmt opsins and opn3 with bovine RH1 showed that both multiple tissue opsins are likely to be functional pigments, consistent with the findings previously reported (Moutsaki et al., 2003, Halford et al., 2001a).

Like many other non-visual opsins (e.g. opn4, opn5, and peropsin), tmt opsins do not possess a negatively charged amino acid at site 113 (numbered as per bovine RH1) (Moutsaki et al., 2003). In visual opsins, the negatively charged residue at site 113 typically serves as a counterion to stabilise the protonated Schiff base linkage between Lys296 and the retinal chromophore (Sakmar, 1998, Sakmar et al., 1989). The presence of a counterion is known to play an important role in the absorption spectrum and photochemical properties of photopigments. By mutating the counterion of bovine RH1 (Glu113) to an uncharged glutamine (Zhukovsky and Oprian, 1989), it causes a blue-shift in the absorption maximum from 500 nm to 380 nm, reflecting deprotonation of the Schiff base linkage. Subsequent replacement of buffer containing a negative halide ion (e.g. chloride, Cl^-) rapidly shifts the absorption peak back towards longer wavelengths (490 nm) (Nathans, 1990). This is because Cl^- ions may act as independent counterions to stabilise protonation from H^+ ions that are already in solution. Since glutamate is found at this position across many vertebrate visual pigments, it has been accepted as the common counterion for the Schiff base in these opsins. However, this is not the case for many non-visual and invertebrate opsins, which carry an uncharged residue at position 113 (Ovchinnikov Yu et al., 1988b, Peirson et al., 2009). With the six zebrafish tmt opsins, their equivalent positions at 113 are all occupied by a neutral tyrosine residue (see Fig. 3.3). It has been suggested that in photopigments like these, their counterions are displaced to a highly conserved Glu181 found in most opsins (Terakita et al., 2004). Exceptions to this are VA opsin, LWS, and MWS pigments, which have been shown

to possess Ser181 (Davies et al., 2010), His181 (Wang et al., 1993), and Tyr181 (Yokoyama et al., 2008) respectively.

It was proposed by Terakita et al. (2004) that in fact Glu181 might have served as an ancestral counterion in ancient vertebrate rhodopsins, predating the role of Glu113. This is perhaps not surprising since the most ancestral opsin class, opn4, also has Tyr113 and Glu181. The difference in counterion position observed in tmt opsins may also provide a clue on the photochemical nature of these pigments. It has been known that monostable vertebrate visual opsins employ Glu113 as a counterion to the Schiff base, whereas many bistable pigments like squid rhodopsin, parapinopsin and melanopsin have their counterions displaced to Glu181 (Tsukamoto and Terakita, 2010, Tsukamoto et al., 2009, Davies et al., 2010, Sekharan et al., 2010). Comparison between two well-characterised opsins, monostable bovine RH1 and bistable squid rhodopsin, demonstrated that they differ in their G-protein activating state during photoactivation. The active state intermediate of RH1 (metarhodopsin II) is found to be UVS and its Schiff base linkage has undergone deprotonation (Emeis et al., 1982), whilst that of the squid rhodopsin (acid metarhodopsin) absorbs visible light and its Schiff base is protonated (Hubbard and Kropf, 1958). Irradiation of the metarhodopsin II thermally converts it into metarhodopsin III, leading to dissociation of the opsin and retinal that defines their monostable characteristic. In contrast, irradiation of squid acid metarhodopsin reverts it back to the original dark state, hence they are bistable pigments (Hubbard and Kropf, 1958). Since tmt opsins are likely to employ Glu181 as their counterion, it was of interest to investigate by UV-vis absorption spectroscopy whether they also exhibit bistability (see Chapter 5).

Bioinformatic analysis showed that multiple *tmt* opsins have been found in different ray-finned fishes (e.g. *Fugu rubripes*), which suggests that ancestors of the teleost lineage may have also possessed copies of the *tmt* opsin gene as a result of whole-genome duplication. It could be that ancient or basal fish have only a few copies of *tmt* opsins, and further duplications might have occurred during the divergence of euteleost lineage. Zebrafish *tmt* opsins can be grouped into pairs as I, II and III subclasses, based on their phylogenetic positions (Fig. 3.3) and greater sequence identities between *tmt* opsin members of the same class compared to those from other subclasses (Table 3.4). These phylogenetic and sequence analyses provide strong evidence that all members of the *tmt* opsin gene family were arisen from a single ancestor through series of gene duplications. The initial duplication event resulted in the formation of *tmtIII* and the ancestor of *tmtI*/*tmtII*. Subsequent duplication of the ancestral *tmtI*/*tmtII* genes produced the ancestors for *tmtI* and *tmtII* groups. Another duplication event in the teleost eventually generated the present-day *tmt* opsins in each of the three classes: *tmt6* and *tmt9* that belongs to the *tmtI* class, *tmt10* and *tmt14* from the *tmtII* class, *tmt2* and *tmt24* from the *tmtIII* class.

In addition, each *tmt* opsin is located on individual chromosomes in zebrafish and the same has been found in medaka (*Oryzias latipes*) (Wayne I.L. Davies, unpublished data). This provided further evidence that supports the hypothesis of whole-genome duplication, which predicted that many of the teleost genes were likely to be copied before the divergence of Ostariophysii (e.g. zebrafish), Protacanthopterygii (e.g. salmon), and Acanthopterygii (e.g. medaka) in the euteleost lineage (Taylor et al., 2001, Meyer and Schartl, 1999, Amores et al., 1998). From

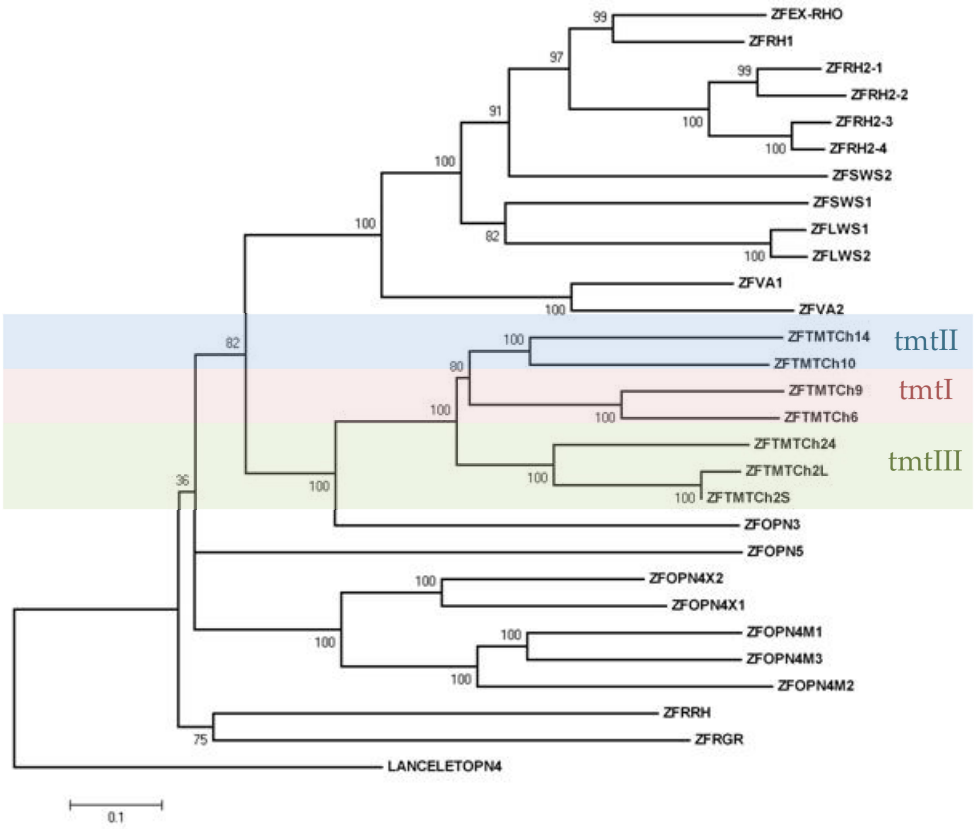
phylogenetic data, it seems that the tmt classes may have initially arose from a single gene into three copies after the divergence of tetrapods from the amniotes. This could be due to: (1) interchromosomal homologous recombination between two chromosomes during meiosis, such that genes are moved from one chromosome to another (Morrow, 2006), or (2) retrotransposition, whereby the coding sequence from the gene is transcribed into mRNA and then into cDNA (Deininger and Batzer, 2002). Genome reinsertion can occur with these cDNA, and sometimes this may replace the original gene (e.g. teleost RH1) (Fitzgibbon et al., 1995) or be inserted randomly as extra copies (e.g. teleost opn4) (Bellingham et al., 2003, Bellingham et al., 2006). Both mechanisms can lead to change in gene location, with the latter mechanism resulting in genes without introns. Since all the *tmt* opsins isolated possess introns, the former mechanism is more likely. In the teleost lineage, the ancestral three copies of tmt may have arisen to become six copies by whole genome duplications (Taylor et al., 2001, Meyer and Schartl, 1999) with subsequent mutations being selected for to result in the six distinct *tmt* opsins identified in zebrafish.

Many duplicated genes are lost over time, but a small number of the gene copies (including *tmt* opsins) were retained in the genome. For genes that have been maintained in the pool, there is a trade off between not gaining mutations to become pseudogenes and being selected for a novel function. It is possible that these additional copies of the *tmt* opsin gene have acquired mutations that could code for new functions, or equally likely, they might have evolved slightly different but complementary functions instead of retaining the original function of the ancestral *tmt* opsin. In both cases, acquisition of new gene functions might allow for ecological

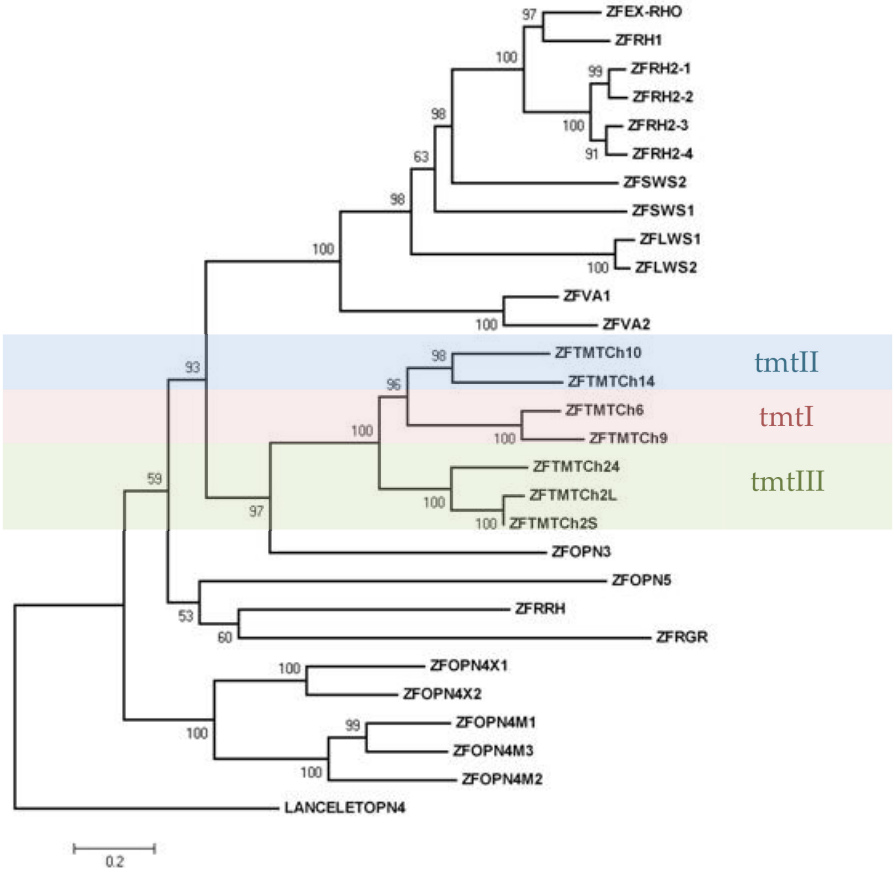
adaptation in an organism that promotes its survival. An example of this has been observed with a different GPCR family, the adrenergic receptors, which consist of five subtypes, α1, α2, β1, β2, and β3 (Yang-Feng et al., 1990). Each of the adrenergic receptor genes have been localised to different chromosome and show distinctive functions in mediating the physiological effects of catecholamines (Small et al., 2003).

Since opsins with similar G-protein coupling and spectral properties often group together phylogenetically (see section 1.3.2; Porter et al., 2007, Shichida and Matsuyama, 2009), the evolutionary relationships of zebrafish *tmt* opsin with other zebrafish photopigments may provide hints on their photochemistry and G-protein activation mechanism. Phylogenetic analyses showed that the full-length nucleotide and putative amino acid sequences of tmt opsins cluster with opn3, which form a sister group to the G_t-coupled opsins (vertebrate visual opsins, VA opsins and exo-rod opsin) (see Fig. 3.4A). Separate phylogenetic trees were also constructed using the sequences of each cytoplasmic loops (CL1, CL2, and CL3) and C-terminal tail of the zebrafish opsins (Fig. 3.4B-E), as these intracellular domains have been implicated for G-protein selectivity and interaction with downstream signalling partners (Scheerer et al., 2008, Yamashita et al., 2000, Kleinau et al., 2010). Interestingly, these data revealed that the intracellular domains of tmt opsins clade together with those of other non-visual opsins (e.g. opn3, opn5, VA opsin) that are closely related to the visual opsins. Given the intermediate phylogenetic positioning of tmt opsins, it is difficult to predict their exact photochemical nature and G-proteins signalling pathways. It is possible that tmt opsins might show similarities to visual as well as non-visual opsins.

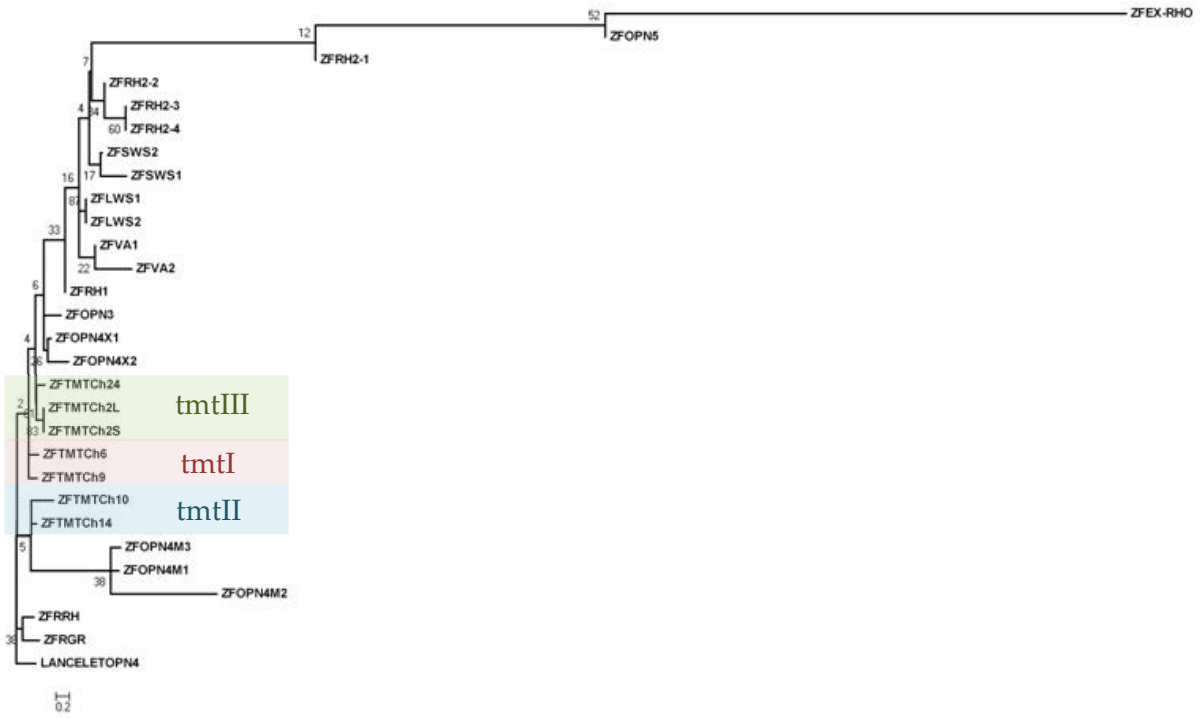
(A- i)


ZFEX-RHO
ZFRH1
ZFRH2-1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFSWS2
ZFSWS1
ZFLWS1
ZFLWS2
ZFVA1
ZFVA2
ZFTMTCh14
ZFTMTCh10
ZFTMTCh9
ZFTMTCh6
ZFTMTCh24
ZFTMTCh2L
ZFTMTCh2S
ZFOPN3
ZFOPN5
ZFOPN4X2
ZFOPN4X1
ZFOPN4M1
ZFOPN4M3
ZFOPN4M2
ZFRRH
ZFRGR
LANCELETOPN4
tmtII
tmtI
tmtIII
0.1

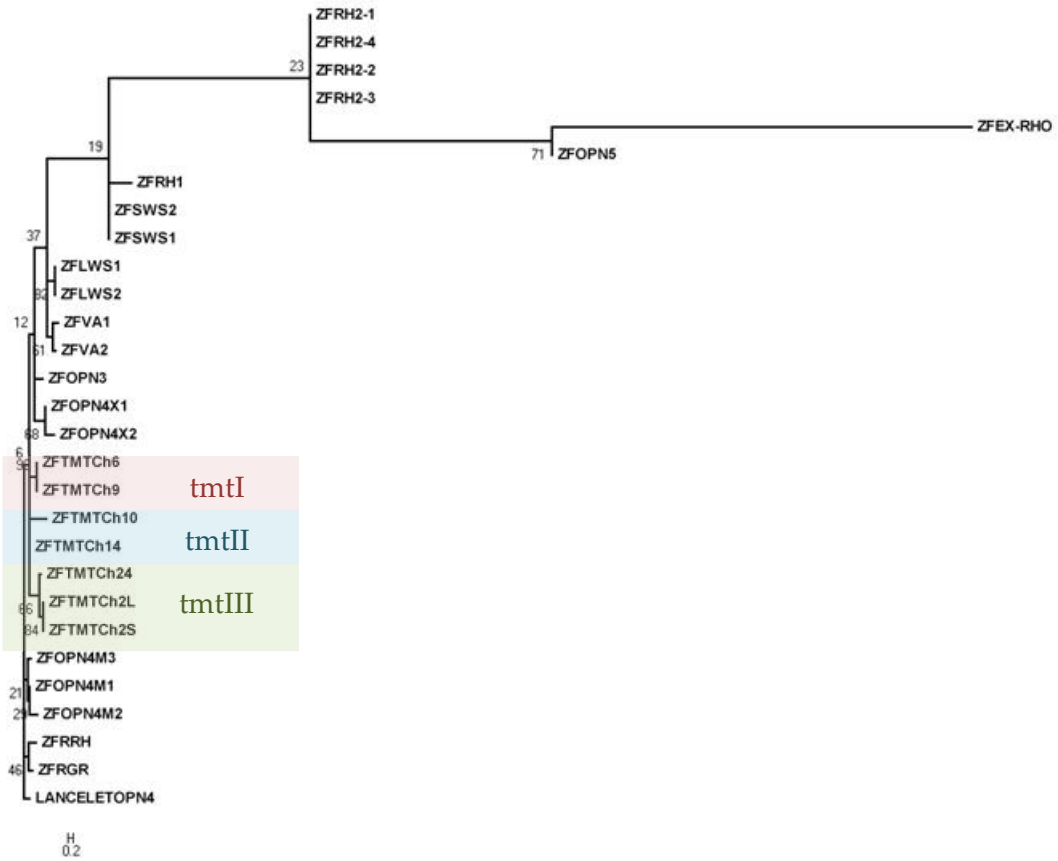
(A- ii)


ZFEX-RHO
ZFRH1
ZFRH2-1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFSWS2
ZFSWS1
ZFLWS1
ZFLWS2
ZFVA1
ZFVA2
ZFTMTCh10
ZFTMTCh14
ZFTMTCh6
ZFTMTCh9
ZFTMTCh24
ZFTMTCh2L
ZFTMTCh2S
ZFOPN3
ZFOPN5
ZFRRH
ZFRGR
ZFOPN4X1
ZFOPN4X2
ZFOPN4M1
ZFOPN4M3
ZFOPN4M2
LANCELETOPN4
tmtII
tmtI
tmtIII
0.2

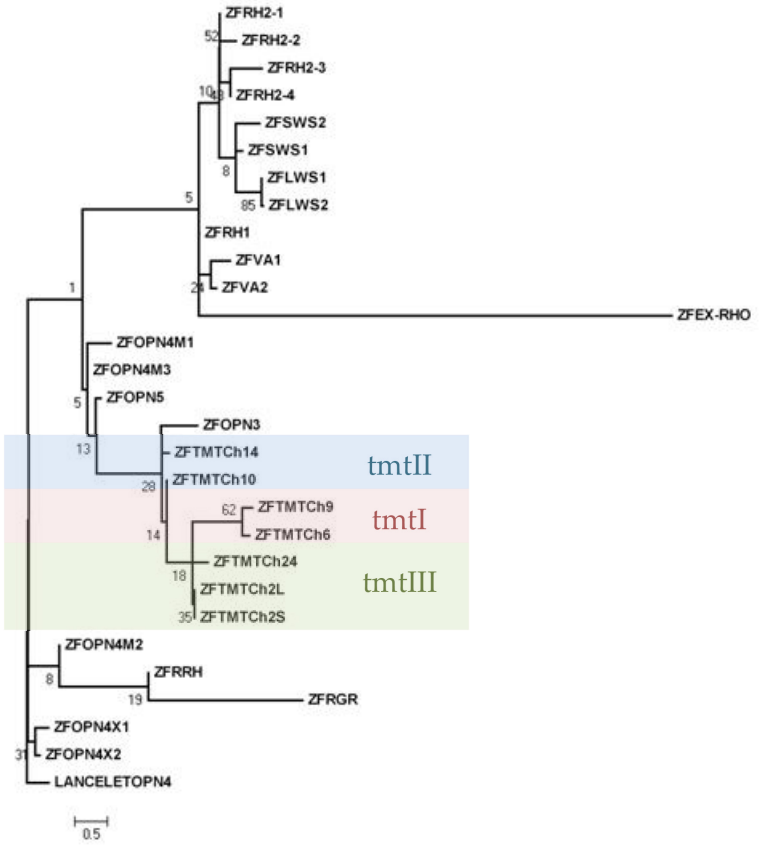
(B-i)


ZFEX-RHO
ZFOPN5
ZFRH2-1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFSWS2
ZFSWS1
ZFLWS1
ZFLWS2
ZFVA1
ZFVA2
ZFRH1
ZFOPN3
ZFOPN4X1
ZFOPN4X2
ZFTMTCh24
ZFTMTCh2L
ZFTMTCh2S
tmtIII
ZFTMTCh6
ZFTMTCh9
tmtI
ZFTMTCh10
ZFTMTCh14
tmtII
ZFOPN4M3
ZFOPN4M1
ZFOPN4M2
ZFRRH
ZFRGR
LANCELETOPN4
0.2

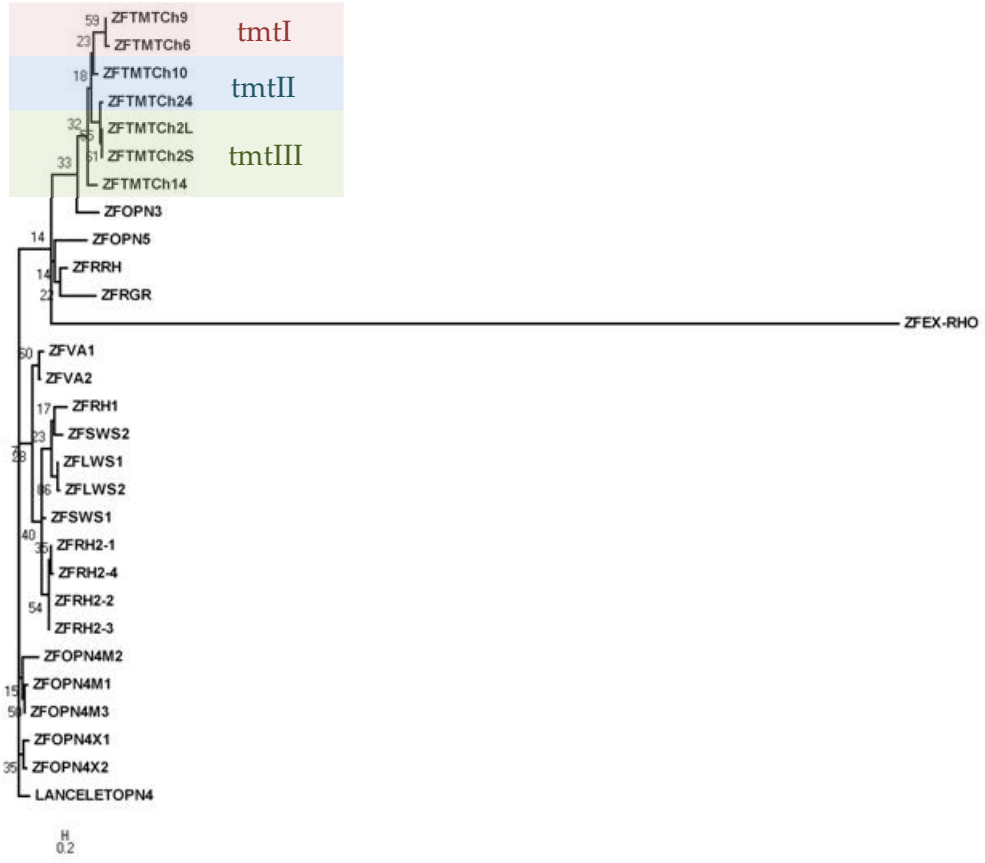
(B-ii)


ZFRH2-1
ZFRH2-4
ZFRH2-2
ZFRH2-3
ZFEX-RHO
ZFOPN5
ZFRH1
ZFSWS2
ZFSWS1
ZFLWS1
ZFLWS2
ZFVA1
ZFVA2
ZFOPN3
ZFOPN4X1
ZFOPN4X2
ZFTMTCh6
ZFTMTCh9
tmtI
ZFTMTCh10
ZFTMTCh14
tmtII
ZFTMTCh24
ZFTMTCh2L
ZFTMTCh2S
tmtIII
ZFOPN4M3
ZFOPN4M1
ZFOPN4M2
ZFRRH
ZFRGR
LANCELETOPN4
0.2

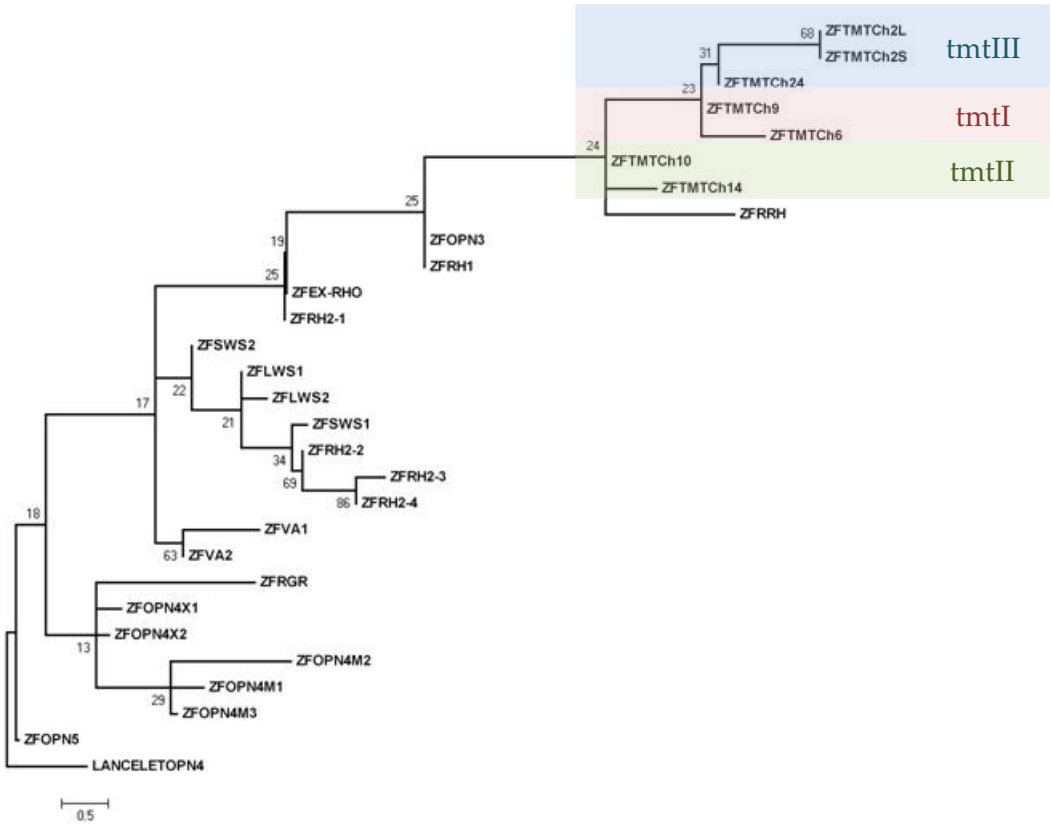
(C-i)


ZFRH2-1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFSWS2
ZFSWS1
ZFLWS1
ZFLWS2
ZFRH1
ZFVA1
ZFVA2
ZFEX-RHO
ZFOPN4M1
ZFOPN4M3
ZFOPN5
ZFOPN3
ZFTMTCh14
ZFTMTCh10
ZFTMTCh9
ZFTMTCh6
ZFTMTCh24
ZFTMTCh2L
ZFTMTCh2S
tmtII
tmtI
tmtIII
ZFOPN4M2
ZFRRH
ZFRGR
ZFOPN4X1
ZFOPN4X2
LANCELETOPN4
0.5

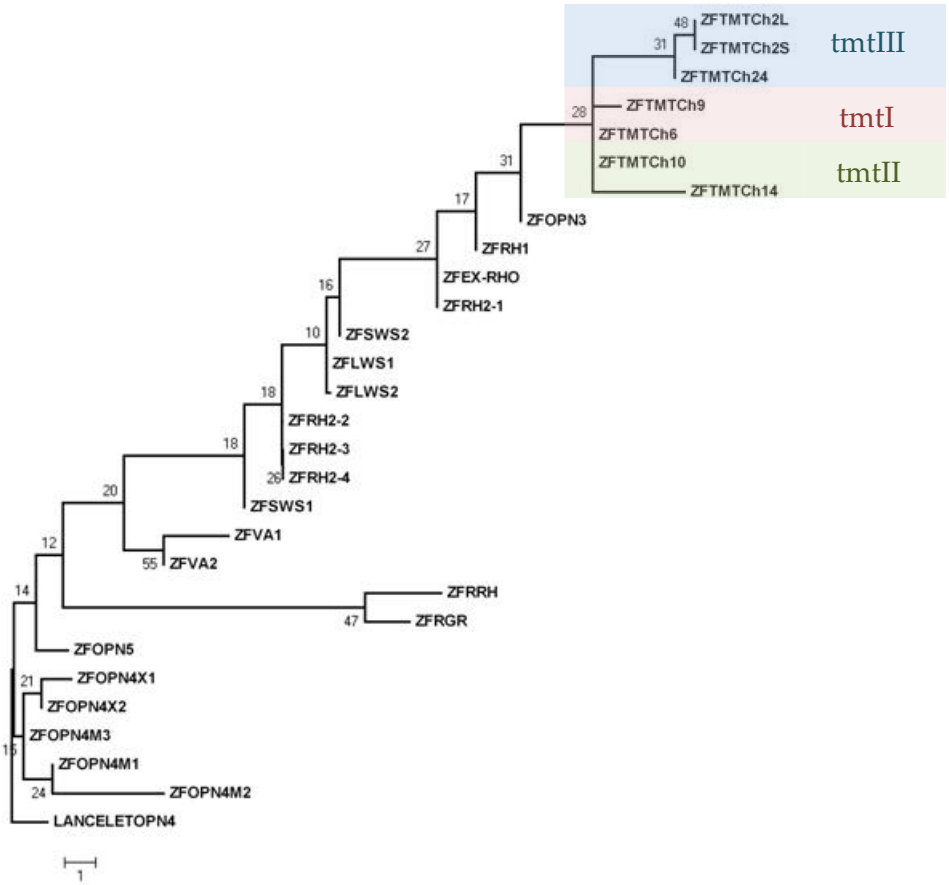
(C-ii)


ZFTMTCh9
ZFTMTCh6
ZFTMTCh10
ZFTMTCh24
ZFTMTCh2L
ZFTMTCh2S
ZFTMTCh14
tmtI
tmtII
tmtIII
ZFOPN3
ZFOPN5
ZFRRH
ZFRGR
ZFEX-RHO
ZFVA1
ZFVA2
ZFRH1
ZFSWS2
ZFLWS1
ZFLWS2
ZFSWS1
ZFRH2-1
ZFRH2-4
ZFRH2-2
ZFRH2-3
ZFOPN4M2
ZFOPN4M1
ZFOPN4M3
ZFOPN4X1
ZFOPN4X2
LANCELETOPN4
0.2

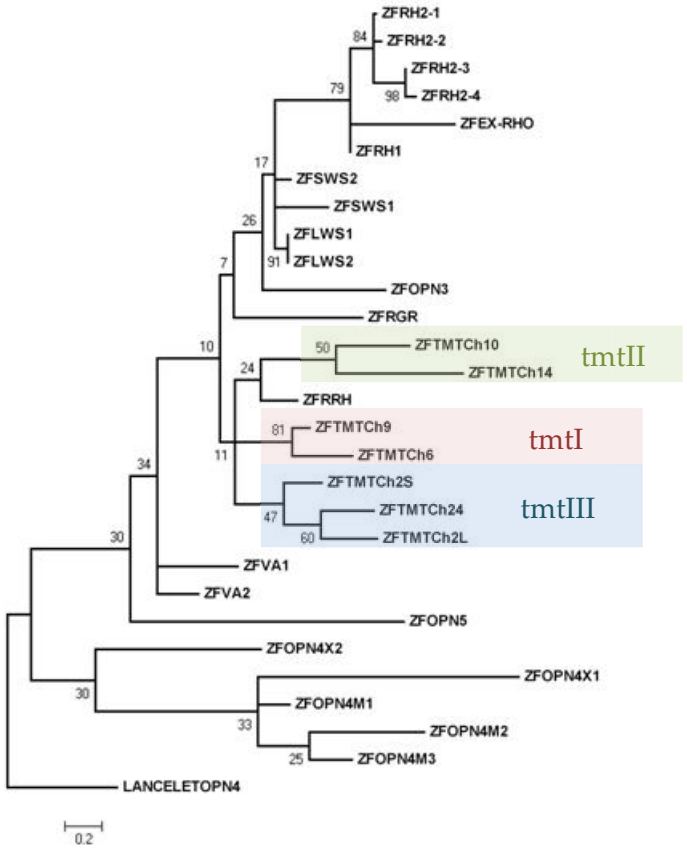
(D-i)


ZFTMTCh2L
ZFTMTCh2S
ZFTMTCh24
ZFTMTCh9
ZFTMTCh6
ZFTMTCh10
ZFTMTCh14
ZFRRH
ZFOPN3
ZFRH1
ZFEX-RHO
ZFRH2-1
ZFSWS2
ZFLWS1
ZFLWS2
ZFSWS1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFVA1
ZFVA2
ZFRGR
ZFOPN4X1
ZFOPN4X2
ZFOPN4M2
ZFOPN4M1
ZFOPN4M3
ZFOPN5
LANCELETOPN4
0.5
tmtIII
tmtI
tmtII

(D-ii)


ZFTMTCh2L
ZFTMTCh2S
ZFTMTCh24
ZFTMTCh9
ZFTMTCh6
ZFTMTCh10
ZFTMTCh14
ZFOPN3
ZFRH1
ZFEX-RHO
ZFRH2-1
ZFSWS2
ZFLWS1
ZFLWS2
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFSWS1
ZFVA1
ZFVA2
ZFRRH
ZFRGR
ZFOPN5
ZFOPN4X1
ZFOPN4X2
ZFOPN4M3
ZFOPN4M1
ZFOPN4M2
LANCELETOPN4
1
tmtIII
tmtI
tmtII

(E-i)


ZFRH2-1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFEX-RHO
ZFRH1
ZFSWS2
ZFSWS1
ZFLWS1
ZFLWS2
ZFOPN3
ZFRGR
ZFTMTCh10
ZFTMTCh14
tmtII
ZFRRH
ZFTMTCh9
ZFTMTCh6
tmtI
ZFTMTCh2S
ZFTMTCh24
ZFTMTCh2L
tmtIII
ZFVA1
ZFVA2
ZFOPN5
ZFOPN4X2
ZFOPN4X1
ZFOPN4M1
ZFOPN4M2
ZFOPN4M3
LANCELETOPN4
0.2

(E-ii)

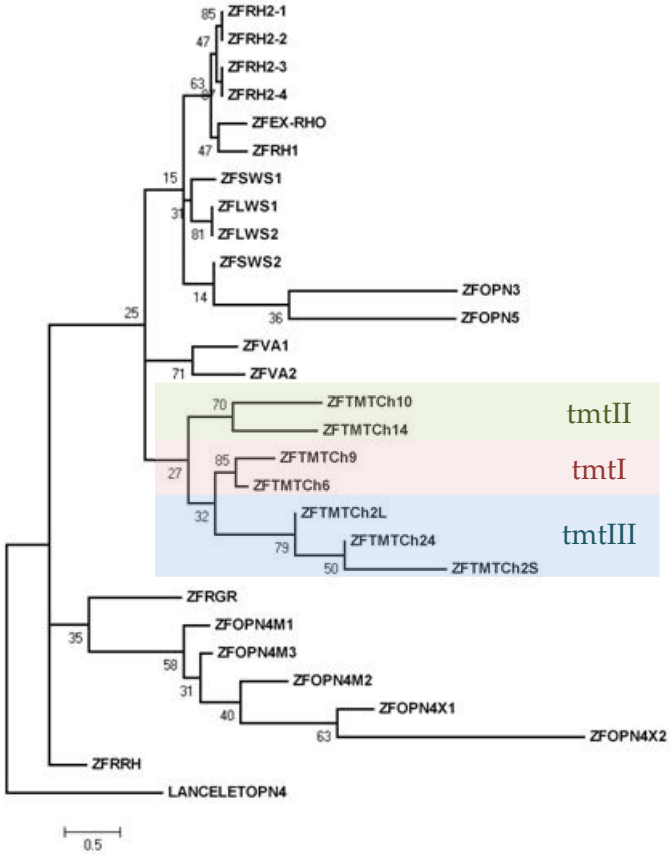

ZFRH2-1
ZFRH2-2
ZFRH2-3
ZFRH2-4
ZFEX-RHO
ZFRH1
ZFSWS1
ZFLWS1
ZFLWS2
ZFSWS2
ZFOPN3
ZFOPN5
ZFVA1
ZFVA2
ZFTMTCh10
ZFTMTCh14
tmtII
ZFTMTCh9
ZFTMTCh6
tmtI
ZFTMTCh2L
ZFTMTCh24
ZFTMTCh2S
tmtIII
ZFRGR
ZFOPN4M1
ZFOPN4M3
ZFOPN4M2
ZFOPN4X1
ZFOPN4X2
ZFRRH
LANCELETOPN4
0.5

Figure 3.4 Phylogeny of all tmt opsins (class I, shaded pink; class II, shaded green; and class III, shaded blue) with other visual and non-visual opsins in *D. rerio* based on their **(A)** full-length coding sequences, cytoplasmic loops **(B)** CL1, **(C)** CL2, **(D)** CL3, and **(E)** C-terminal domain (Wayne, I.L. Davies, unpublished data). The phylogenetic trees were constructed on codon-matched sequence alignments of **(i)** the nucleotide sequences using Maximum Composite Likelihood (MCL) (Tamura and Nei, 1993),and **(ii)** the amino acid sequences using Jones-Taylor-Thornton (JTT) substitution matrix (Jones et al., 1992). Neighbour-joining trees were generated with 1000 bootstrapping replicates. Internal branch supports are expressed as percentages at the node of each branch. The scale bars at the bottom of figures denote the number of nucleotide or amino acid substitutions per site. See Section 2.3.6 for detailed methodology and the sequences used.

Evidence for the existence of peripheral clocks in zebrafish was first demonstrated by Whitmore et al. (1998). Their findings showed that these clocks exhibit robust circadian oscillation in a wide range of tissues outside the CNS, independent from the master clock in the pineal gland. Moreover, the rhythmicity observed can be synchronised to the daily cycle by light (photoentrainment). Thus, it was concluded that there must be photoreceptive elements widely present across all tissues in the zebrafish (Whitmore et al., 2000, Whitmore et al., 1998). Almost twenty years later, strong evidence for not only one, but six different tmt opsins have been identified as potential "peripheral photopigments" in zebrafish. At present, it is difficult to account for the multiplicity of this opsin subfamily and there could be various explanations including whole-genome and independent duplication events. Nonetheless, functional data in the following chapters will provide important clues to the diversity observed in this group. Taken together, the molecular architectures of tmt opsins, their remarkably conserved genomic structures across different classes (including critical residues Lys296, Glu181, Cys118, Cys187, E(D)RY, and NPxxY$(x)_{5,6}$F motifs) and their similarities in amino acid sequences to other known opsins, have all so far pointed to a shared evolutionary history for these pigments.

CHAPTER 4

Temporal and Spatial Expression of Zebrafish tmt Opsins

CHAPTER 4: Temporal and Spatial Expression of Zebrafish tmt Opsins

4.1 Introduction

When tmt opsin was first discovered in fugu (*Fugu rubripes*) and zebrafish (*Danio rerio*), one of the most intriguing findings was that its widespread expression profile was consistent with that of the circadian clock genes in teleosts (Moutsaki et al., 2003). This opsin was detected in nearly all tissues studied, including eye, brain, kidney, heart, and even clock-containing embryonic cell-lines (PAC2 and Z3) of zebrafish. Importantly, studies have shown that not only do these organs and cell-lines display intrinsic circadian oscillations in clock gene expression, but that these rhythms are also photoentrainable (Stokkan et al., 2001, Pando et al., 2001, Sakamoto et al., 2004, Wu et al., 2010). The deduced amino acid sequence of this tmt opsin possesses all the essential features of a classical photopigment (e.g. K296 for retinal binding, C110 and C187 for conformational stability). Since the putative structure of tmt opsin and its expression pattern coincide with CNS as well as peripheral circadian clocks in zebrafish, it has been proposed as a potential photopigment (Moutsaki et al., 2003), one which may have a role in mediating the entrainment of circadian rhythms. Evidence for this was supported by a recent study showing that mutated *tmt* opsin and *opn4* genes were found in a species of Somalian cavefish (*Phreatichthys andruzzii*) which have no entrainable clocks, suggesting a role for these two genes in peripheral entrainment (Cavallari et al., 2011).

RT-PCR experiments detailed in Chapter 3 showed that there are in fact at least six different *tmt* opsins in zebrafish. It would be important to know if the

expression of these opsin genes correlates with the developmental time course of circadian clocks in the teleost, which was shown to be functional from day one of embryogenesis (Dekens and Whitmore, 2008). This present study aimed to discover if these six *tmt* opsins have overlapping pattern of localisation with distinct functional roles in developing embryos of zebrafish and in their adult tissues.

To understand the molecular mechanisms governing the expressing pattern of *tmt* opsins, it is necessary to identify and examine their promoters. At the transcriptional level, the production of mRNA is controlled by the conformation of DNA and its interactions with various transcription factors, enhancers and repressors (Pesole et al., 2000). Binding of these regulatory factors is mainly based on their specificity for consensus sequences within the promoter region upstream of the transcription start site, e.g. TATA binding protein (TBP) binding to nucleotide sequence 5′-TATAWAWN-3′ (Juo et al., 1996). Expression of tmt opsins can also be modulated at the level of translation, which involves production of tmt opsin proteins from the mRNA transcripts. Translation is primarily regulated at the initiation step, when ribosomal subunits are recruited. Secondary structure of mRNA (e.g. hairpin loops; Gingras et al., 1999) and the presence of protein binding sites (e.g. Kozak consensus sequence, GCCRCCAUGG; Kozak, 1991b) flanking the translational start codon (AUG) are important contributing factors that can affect the binding of ribosomes and thus control the efficiency of the translation process.

The presence of multiple tmt opsins in zebrafish may suggest they have been selected for on the basis of spatial, temporal or functional diversity. Studying the spatial-temporal expression pattern of each *tmt* opsin gene may, therefore, provide

clues on their functional roles. Given the potential association of *tmt* opsins with the photoentrainment of peripheral circadian system observed in teleosts (Moutsaki et al., 2003), this chapter examined the expression pattern of all six *tmt* opsin genes in developing zebrafish embryos by whole-mount RNA *in situ* hybridisation (RNA ISH). Since the eye and the brain of adult zebrafish are well-established light-sensitive organs that exhibit pronounced circadian rhythms (Mueller and Neuhauss, 2012, Vigh et al., 2002, Menger et al., 2005), these two regions were also investigated by RNA ISH. In addition, the promoter region of *tmt* opsins were analysed for regulatory mechanisms that may control the expression patterns observed.

4.2 Results

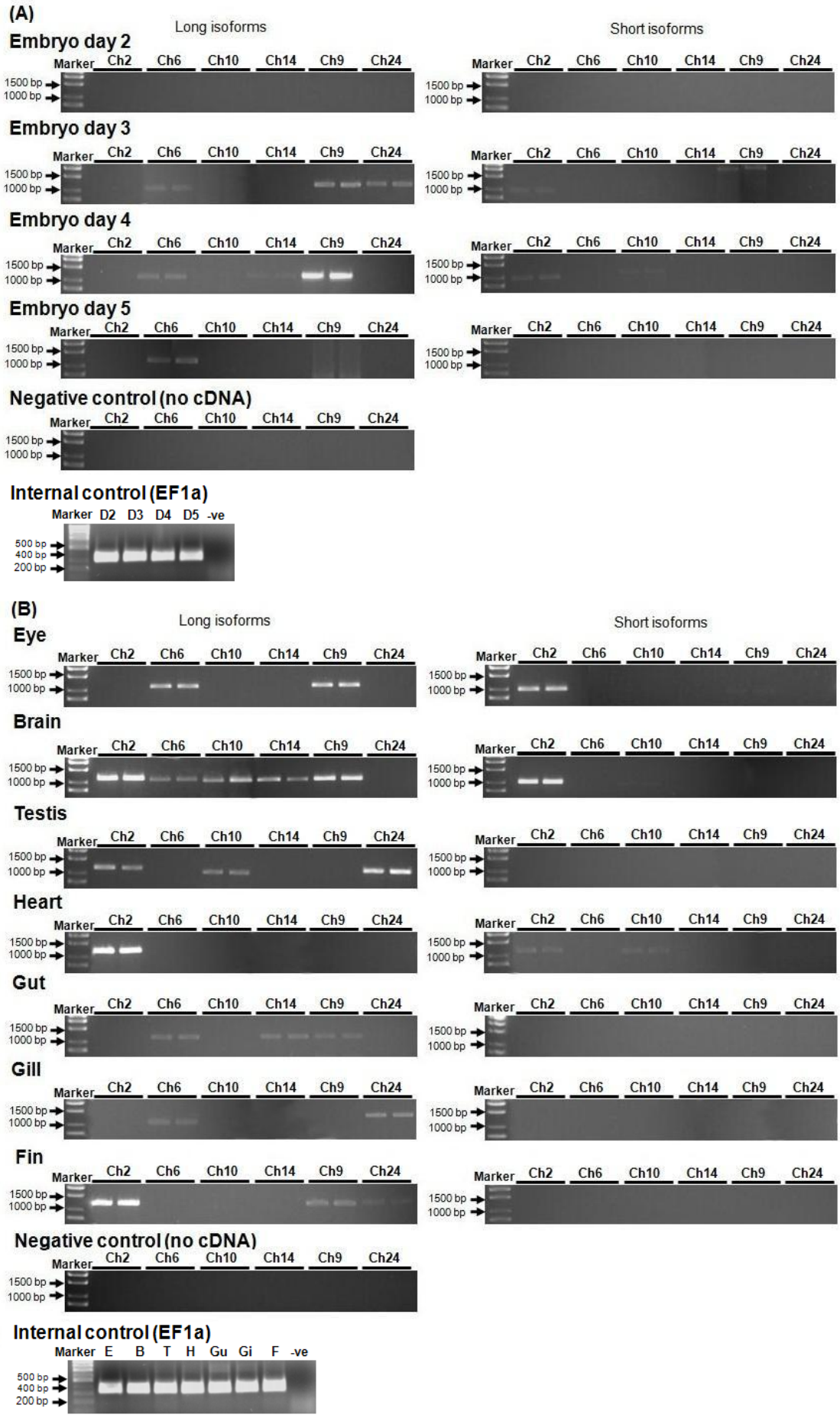
4.2.1 RT-PCR of *tmt* opsins in developing and adult zebrafish

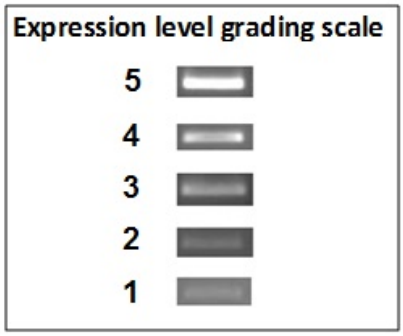
To study the relative expression pattern of *tmt* opsins during the early developmental period of zebrafish and in their adult tissues, reverse transcription-polymerase chain reactions (RT-PCRs) of *tmt* opsin gene transcripts were performed in day 2-5 embryos and in a panel of seven zebrafish adult tissues (eye, brain, testis, heart, gut, gill, fin). These tissues have all been shown to contain light-entrainable clock gene expression, suggesting the presence of photopigment that may contribute to light detection (Whitmore et al., 1998, Whitmore et al., 2000, Dekens et al., 2003, Kaneko et al., 2006). Although zebrafish embryos are known to exhibit photoentrainable clock gene expression from day 1 of embryogenesis, day 1 embryos were not available at the time of experiment. Expression of *tmt* opsins in day 1 embryos were later examined by whole-mount ISH and nCounter® Analysis (NanoString Technologies).

Previous RT-PCR data showed that *tmt2* has two isoform variants, *tmt2l* and *tmt2s*. The long isoform (*tmt2l*) has four exons, whilst the short isoform (*tmt2s*) has three exons with the third exon extending into the next available stop codon within intron 3 (Section 3.2.1). Based on this, predictions were made about the putative short variants of other tmt opsins, with reverse primers designed to the first stop codon located in intron 3 to detect their presence and absence. The primers used for isolating long and short isoforms of *tmt* opsins are listed in Table 2.1.

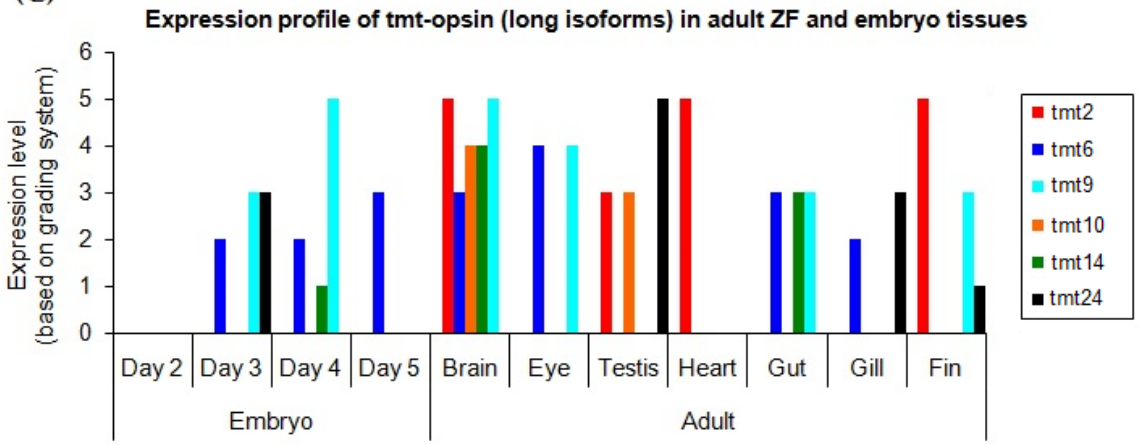
From the RT-PCR data shown in Fig. 4.1A and C, not all of the *tmt* opsins were detected at the same time in zebrafish embryos. It appears from the results that no *tmt* opsins transcripts were detected in day 2 embryos. From day 3 onwards, the short isoform *tmt2s* and long isoforms of *tmt6*, *tmt9*, and *tmt24* were detected. In adult zebrafish, *tmt* opsin expression varied between different tissues, as well as within the classes (Fig. 4.1B, C). Brain expressed most of the *tmt* opsins, which included long isoforms of *tmt2*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, and the short isoforms *tmt2s* and *tmt10s*. All the other tissues tested expressed three to four of the *tmt* opsins at varying levels, but none of the *tmt* opsins were restricted to a specific tissue. Overall, the expression of long isoforms seemed generally high (Fig. 4.1C), whilst that of the short isoforms was variable (Fig. 4.1D). Differential expression of the isoform variants suggests that they could be regulated by a general mechanism controlling their expression in different tissues as well as during the development of zebrafish, which will be discussed later.

Since RT-PCR is not generally quantitative, the relative expression levels of *tmt* opsins were confirmed by NanoString Analysis in Section 4.2.5.


(A)
Long isoforms
Short isoforms
Embryo day 2
Embryo day 3
Embryo day 4
Embryo day 5
Negative control (no cDNA)
Internal control (EF1a)
Marker Ch2 Ch6 Ch10 Ch14 Ch9 Ch24
1500 bp
1000 bp
Marker D2 D3 D4 D5 -ve
500 bp
400 bp
200 bp
(B)
Long isoforms
Short isoforms
Eye
Brain
Testis
Heart
Gut
Gill
Fin
Negative control (no cDNA)
Internal control (EF1a)
Marker E B T H Gu Gi F -ve


Expression level grading scale
5
4
3
2
1

(C)


Expression profile of tmt-opsin (long isoforms) in adult ZF and embryo tissues
Expression level (based on grading system)
0
1
2
3
4
5
6
Day 2
Day 3
Day 4
Day 5
Brain
Eye
Testis
Heart
Gut
Gill
Fin
Embryo
Adult
tmt2
tmt6
tmt9
tmt10
tmt14
tmt24

(D)

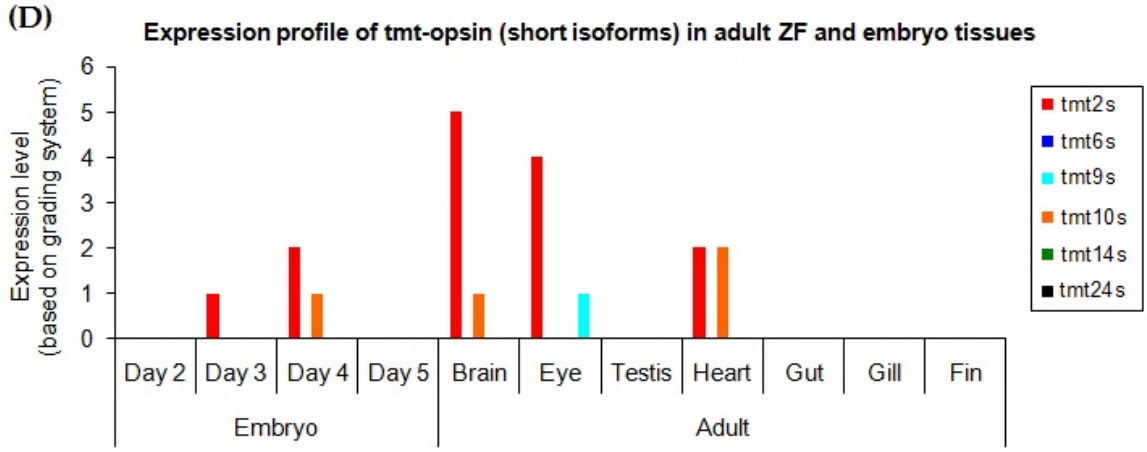

Expression profile of tmt-opsin (short isoforms) in adult ZF and embryo tissues
Expression level (based on grading system)
0
1
2
3
4
5
6
Day 2
Day 3
Day 4
Day 5
Brain
Eye
Testis
Heart
Gut
Gill
Fin
Embryo
Adult
tmt2s
tmt6s
tmt9s
tmt10s
tmt14s
tmt24s

Figure 4.1 RT-PCR analyses of *tmt* opsin mRNA extracted at different stages of zebrafish embryonic development and in a panel of adult tissues. **(A)** PCRs performed on zebrafish embryo tissues day 2-5 results in the amplification of coding regions from four long *tmt* opsin variants (*tmt2*, 1187 bp; *tmt9*, 1169 bp; *tmt14*, 1163 bp; *tmt24*, 1217 bp) and two of the predicted short isoforms (*tmt2*, 989 bp; *tmt10*, 1019 bp) in tissues from day 3 onwards. Each reaction sample was loaded over two adjacent wells. **(B)** PCRs performed on adult zebrafish eye, brain, testis, heart, gut, gill and fin results in the amplification of the coding regions from all long variants of six *tmt* opsins (*tmt2*; *tmt6*, 1136 bp; *tmt10*, 1097 bp; *tmt14*; *tmt9*; *tmt24*) and three of the short isoforms (*tmt2*, *tmt10*, and possibly *tmt9* (995 bp)) from different tissues. Negative controls for both (A) and (B) (no cDNA) were included, with amplification of the elongation factor 1-alpha (EF1a) being used as an internal positive control. These data indicate that similar to *tmt2*, long and short isoforms of *tmt9* and *tmt10* are likely to exist. However, the short isoforms of *tmt9* and *tmt10* seem to be present at very low levels. **(C)** Graph showing the qualitative expression profile of full-length long isoform *tmt* opsins in embryonic and adult zebrafish tissues. Expression levels were graded subjectively, based on the scale shown above the diagrams, and plotted in graphic form. **(D)** Graph showing the expression profile of the full-length short isoforms of *tmt* opsins in adult zebrafish and embryo tissues. Data were recorded in the same manner as graph (C).

4.2.2 Spatial expression patterns of *tmt* opsins during early embryonic development of zebrafish

The spatio-temporal expression of the six *tmt* opsin genes were analysed in day 1 to day 5 zebrafish embryos by whole-mount RNA ISH (Fig. 4.2). The mRNA of *tmt2* was shown to be expressed earliest in development, with transcripts being detected in the notochord (NC) and neural tube (NT) at day 2 (Fig. 4.2A). Although RT-PCR showed no amplification of *tmt* opsin mRNA from day 2 embryo (Section 4.2.1), this could be due to the low abundance of transcripts present in the sample. In day 3 embryos, RNA ISH data showed that *tmt6*, *tmt9*, *tmt10*, and *tmt24* also began to be expressed in the notochord (Fig. 4.4F, J, N and V), whilst *tmt2* was observed in the neurocranial trabecula region (T) (Fig. 4.2B). Similar data was obtained from RT-PCR, which confirmed the expression of *tmt6*, *tmt9* and *tmt24* in day 3 embryos, but not *tmt2* and *tmt10* (probably also due to low expression level). In day 4 and day 5 embryos, all tmt opsins were detected by RNA ISH in the notochord. In addition, positive signals were observed in the gut (G) and liver (L) for all the tmt opsins and in the medulla oblongata (M) for *tmt2*, *tmt6*, *tmt9* and *tmt10* (Fig. 4.2D, H, L and P). However, RT-PCR analysis of day 4 and day 5 embryos showed that only *tmt6*, *tmt9*, and *tmt14* transcripts were amplified. The inconsistency in the RT-PCR data could be due to the quality of the cDNA samples tested. Overall, it appears from RNA ISH data that *tmt* opsins are expressed very early in development from day 2.

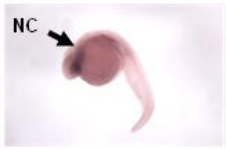
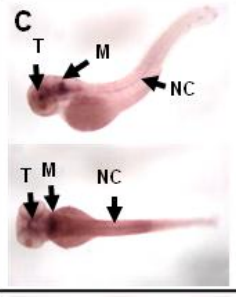
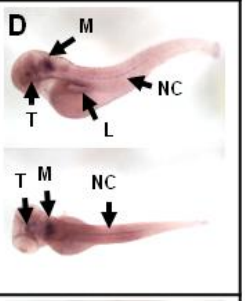
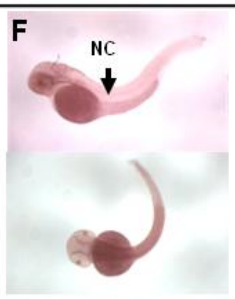
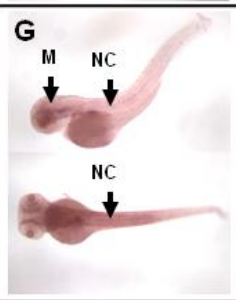
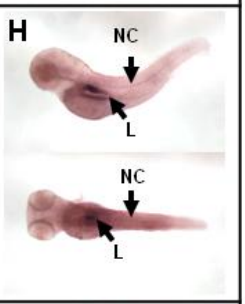
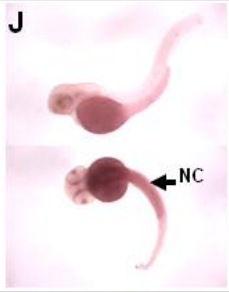
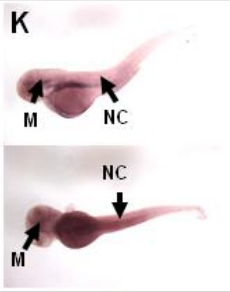
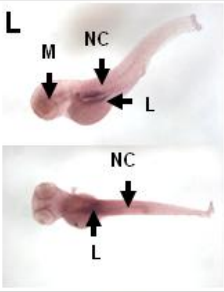
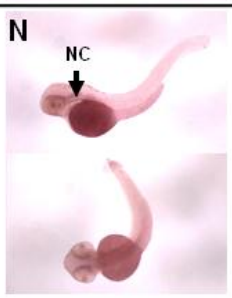
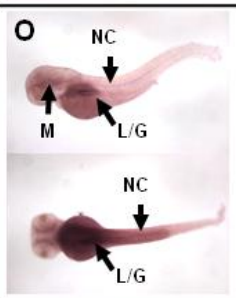
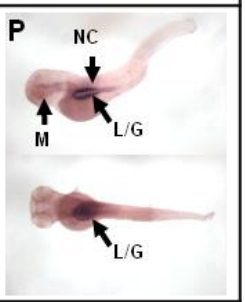
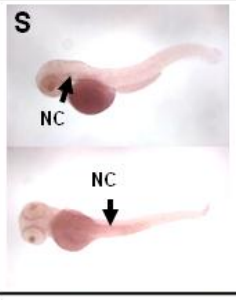
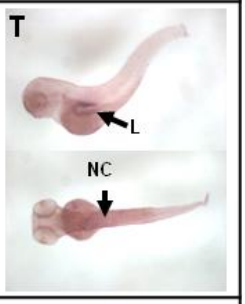
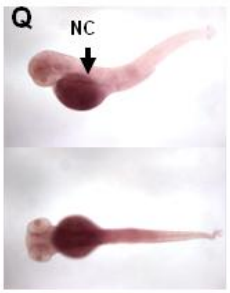
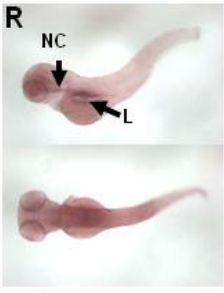
	Day 2	Day 3	Day 4	Day 5
tmt2	A 	B 	C 	D 
tmt6	E No expression observed	F 	G 	H 
tmt9	I No expression observed	J 	K 	L 
tmt10	M No expression observed	N 	O 	P 
tmt14	Q No expression observed	R 	S 	T 
tmt24	U No expression observed	V 	Q 	R 

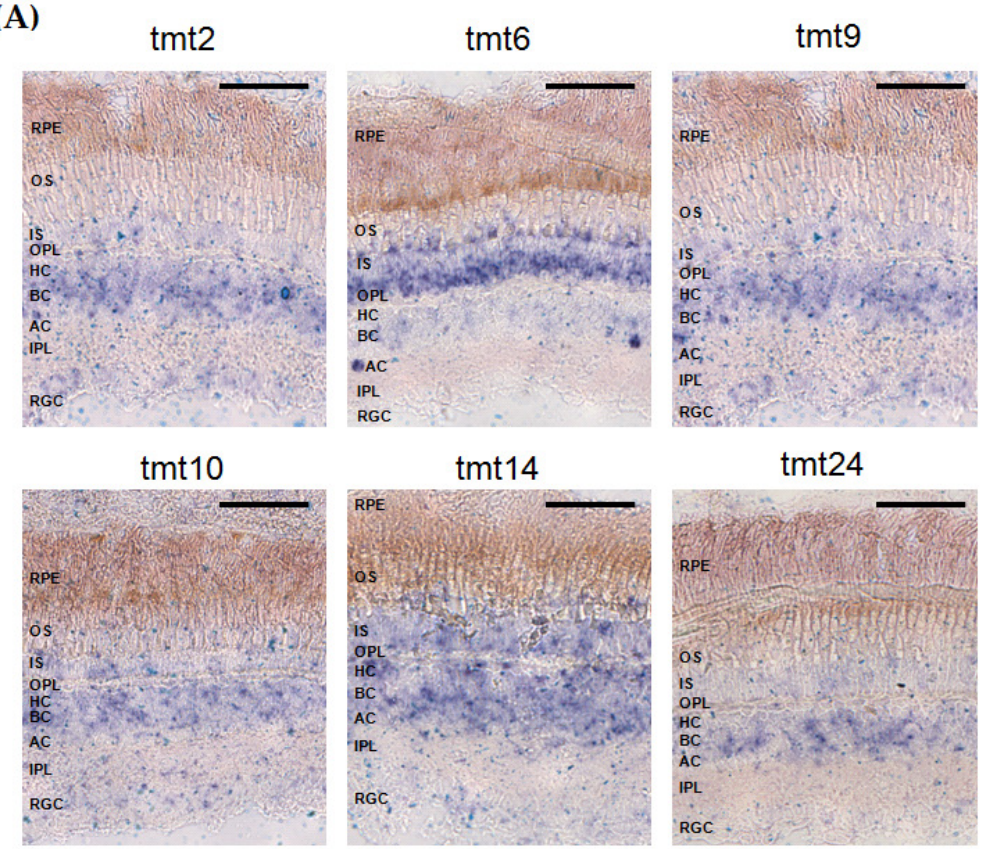
Figure 4.2 The expression pattern of long six *tmt* opsin transcripts: *tmt2* **(A-D)**, *tmt6* **(E-H)**, *tmt9* **(I-L)**, *tmt10* **(M-P)**, *tmt14* **(Q-T)** and *tmt24* **(U-R)** during the development of zebrafish embryos. Embryos from day 1-5 were collected and hybridised with antisense tmt opsin riboprobes using whole mount in situ hybridisation. Data from day 1 embryos were not shown here, as there was no signal was observed. For C, D, F, G, H, J, K, L, N, O, P, S, T, V, Q, and R, the top panels showed a dorsal view, whilst the bottom panels were lateral views. The arrows indicate where strong tmt riboprobe signals were observed: NC, notochord; T, neurocranial trabecula; M, medulla oblongata; G, gut; L, liver.

4.2.3 Spatial expression patterns of tmt opsin in the retina of adult zebrafish

To determine the localisation pattern of *tmt* opsins within the zebrafish eye, RNA ISH was performed on adult retina sections. The results revealed that all six *tmt* opsins were detected in the eye with overlapping profiles of expression (Fig. 4.3A, B).

Transcripts of *tmt* opsins were detected predominantly in the inner nuclear layer, with bipolar cells (BCs) strongly expressing *tmt2*, *tmt9*, *tmt10*, and *tmt14*, and to a lesser extend *tmt6* and *tmt24* (Fig. 4.3A). A closer observation of the inner retina at the level of the BCs showed that not all cells were labelled but signals were observed at regular intervals. For example, the mRNA signals of *tmt2* and *tmt9* were more evenly distributed amongst the BCs as compared to those of *tmt6* and *tmt24*. This indicates that each tmt opsin has a distinctive expression pattern in different subpopulations of BCs, and it is likely that some of these transcripts are co-expressed in the same cell types. A few single cells located close to the inner plexiform layer (IPL), resembling a subset of amacrine cells (ACs), were also labelled at regular interval, showing strong expression of *tmt6* mRNA. In retinal ganglion cells (RGCs), low levels of expression were observed for *tmt2*, *tmt9*, *tmt10*, and *tmt14*, and no mRNA was detected for *tmt6* and *tmt24*. In the retinal pigment epithelium (RPE), only *tmt9* and *tmt10* were detected at low levels. Interestingly, all *tmt* opsin transcripts were detected in the inner segment (IS) layer containing cell bodies of the photoreceptors. It is important to note that the precise cellular identification of *tmt* opsin expression in this study may be complicated by their potential expression in the Müller cell fibres, which span across the whole retina.

(A)



(B)

	tmt2	tmt6	tmt9	tmt10	tmt14	tmt24
RPE	-	-	++ (chloroid)	++ (chloroid)	-	-
OS	-	-	-	-	-	-
IS (rods and/or cones	++	++++	++	+++	++	+
HC	-	-	-	-	-	-
BP	++++	++	+++	+++	+++	+
AC	-	++++ (subtype)	-	-	-	-
RGC	++	-	+	+	+	-

Figure 4.3 Expression of six *tmt* opsin genes (*tmt2*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, and *tmt24*) in the adult eyes of zebrafish. **(A)** RNA *in situ* hybridisation. Scale bars, 50 μm **(B)** Table illustrating global expression pattern of the *tmt* opsins in different cell types of the retina. Retinal pigment epithelium, RPE; outer segment, OS; inner segment, IS; outer plexiform layer, OPL; horizontal cell, HC; bipolar cell, BC; amacrine cell, AC; inner plexiform layer, IPL; retinal ganglion cell, RGC.

Overall, the expression of *tmt* opsins seems to extend across the entire retina and is found in most neuronal cell types apart from horizontal cells (see Fig. 4.3B). RNA ISH revealed the presence of all *tmt* opsins in the eyes as compared to only *tmt6* and *tmt9* detected by RT-PCR analysis, indicating that RNA ISH is a more sensitive technique than DNA detection. It is possible that mRNAs of *tmt* opsins are also present in the eyes of the zebrafish embryos. However, their expression levels were too low to be determined visually from whole-mount ISH.

4.2.4 Expression of tmt opsin in the brain of adult zebrafish

RNA ISH was also performed on adult zebrafish brain sections to determine the expression pattern of *tmt* opsin transcripts. It was found that *tmt* opsin RNAs were expressed in various locations of the brain, which is summarized in Table 4.1.

In the forebrain, *tmt6* and *tmt9* were detected in the external cellular layer (ECL) of the olfactory bulb and in the lateral olfactory tract (LOT) (Fig. 4.5A, 4.6A & B). Throughout the telencephalon, the expression of different *tmt* opsins was distributed in different regions of the ventral, dorsal, and medial areas. All six *tmt* transcripts were detected along the telencephalic ventricle in the dorsal (Vd), ventral (Vv), and postcommissural nuclei of the ventricle (Vp) (Fig. 4.4A, 4.5B, 4.6A, 4.7A, 4.8A, 4.9A). In the dorsal telencephalic area, *tmt9*, *tmt10* and *tmt14* were expressed in the medial zone of D (Dm) (Fig. 4.6B, 4.7B, 4.8B). Additionally, *tmt9* was detected in the supracommisural nuclei of the ventricle (Vs), with weak labelling observed in the dorsal (ENd), as well as ventral (ENv), parts of the entopeduncular nuclei (Fig. 4.6B).

Riboprobe signals were also observed in localised areas of the midbrain. Within the preoptic area, *tmt6* and *tmt14* were weakly detected in the ventral havenula (Hav) (Fig. 4.5C) and the anterior part of the parvocellular preoptic nucleus (Ppp) (Fig. 4.8A), respectively. Towards the pretectum of the midbrain, *tmt2* and *tmt9* transcripts were detected in the dorsal (PPd) and ventral (PPv) portions of the periventricular prectectal nucleus (Fig. 4.4B, 4.6C). In addition, *tmt9* appeared to be present in the commissural posterior (Cpost) and in the periventricular nucleus of the posterior tuberculum (Tpp) (Fig. 4.6C). In the most ventral part of the midbrain, *tmt10* was faintly labelled in the dorsal zone of periventricular hypothalamus (Hd), the medial periglomerular nuclei (PGm), the tertiary gustatory nucleus (TGN), and the posterior tuberal nuclei (PTN) (Fig. 4.7D). Towards the most caudal part of the midbrain, all six *tmt* opsins showed intense staining in the periventricular gray zone of the optic tectum (PGZ) (Fig. 4.4B, 4.5D, 4.6C, 4.7D, 4.8C, 4.9B). Apart from *tmt2*, *tmt6* and *tmt10*, the other three *tmt* opsins were detected in the torus longitudinalis (TL) (4.6C, 4.8C, 4.9B). There was also a lower level of staining observed for *tmt9* in the nucleus of the medial longitudinal fascicle (Nmlf) (Fig. 4.6D).

In the rhobencephalon, successive sections showed expression of all six *tmt* opsins in the corpus cerebella (CCe) and the lobus caudalis cerebella (LCa). In these regions, expression of *tmt2*, *tmt9* and *tmt10* (Fig. 4.4D & E, 4.6G & I, 4.7F & G) were more intense than those of *tmt6*, *tmt14* and *tmt24* (Fig. 4.5E & F, 4.8H & I, 4.9 E-H). In a region adjacent to the cerebella caudal lobe, corresponding to the eminentia granularis (EG), expression of *tmt2*, *tmt9* and *tmt10* were strongly detected (Fig. 4.4D, 4.6G, 4.7F). While neither *tmt14* nor *tmt24* were found in the medulla oblongata,

other *tmt* opsins were expressed in this region. Hybridisation signals were detected with high intensity in the outermost layer surrounding the lobus facialis (LVII) and the lobus vagus (LX) for *tmt2*, *tmt6*, *tmt 9* and *tmt10* (Fig. 4.4G & H, 4.5 G& H, 4.6J, 4.7I), in the comissura infima of Haller (Cinf) for *tmt9*, and in the crista cerebellaris (CC) for *tmt10* (Fig. 4.7 H).

In general, these data showed that there is a degree of overlapping gene expression amongst the six *tmt* opsins. Since the RNA ISH experiments were performed using full-length probes of the paralogous genes sharing large regions of similarity, it is important to note that there is a potential problem associated with cross-reactivity. By designing shorter probes targeting unique regions of the genes in future experiments, this may enhance their binding specificity.

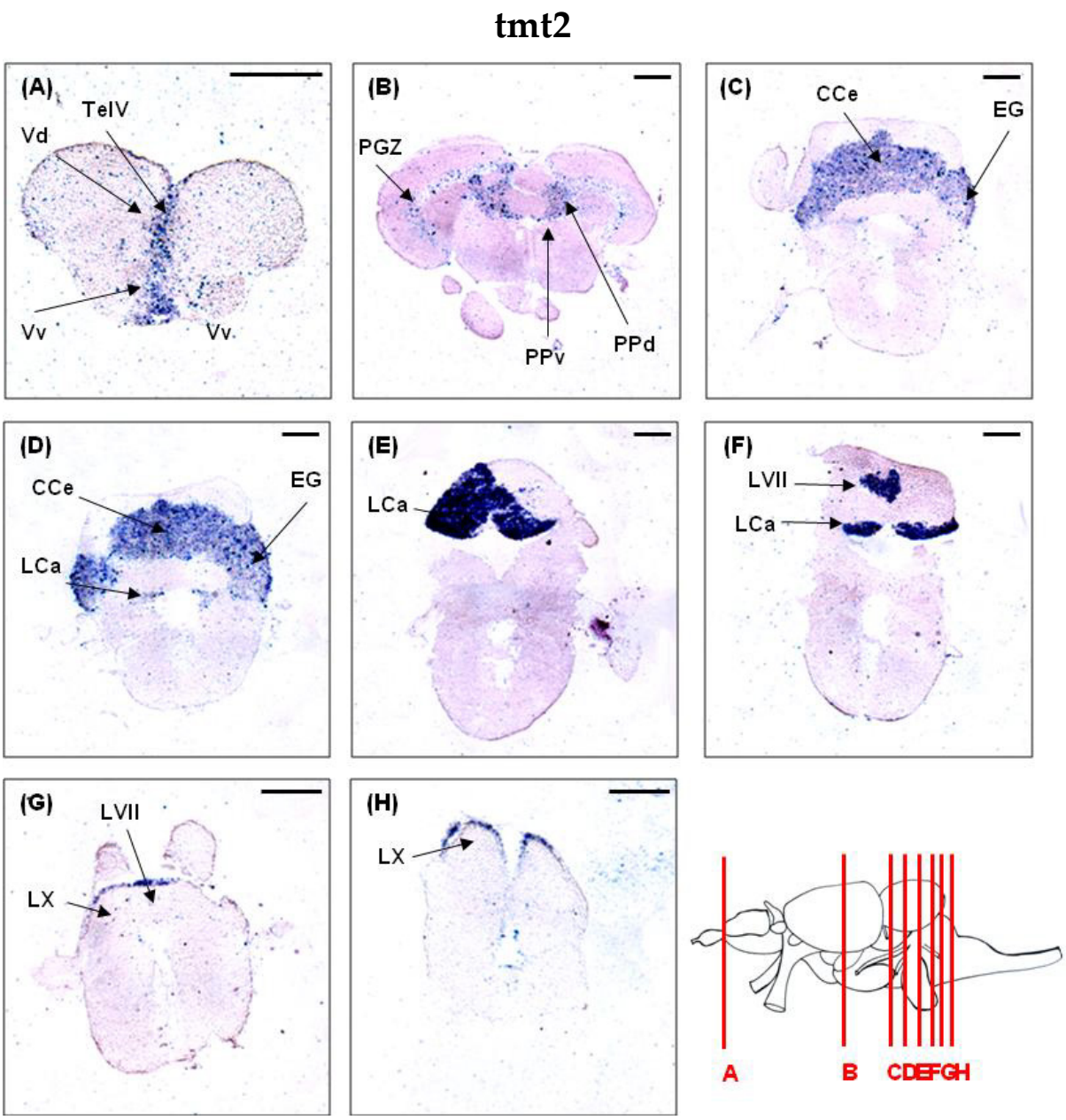


Figure 4.4 Expression of the *tmt2* gene in the adult zebrafish brain. *In situ* hybridisation detection of *tmt2* transcripts was performed on transverse rostro-caudal sections (A-H). *tmt2* expression was detected in: **(A)** the telencephalic ventricle (TelV), dorsal (Vd), ventral (Vv), and postcommisural (Vp) nuclei of the ventral telencephalon; **(B)** the dorsal (PPd) and ventral (PPv) parts of the periventricular prectectal nuclei, and the periventricular gray zone (PGZ) in the mesencephalon; **(C-E)** corpus cerebellum (CCe), eminentia granularis (EG), and lobus caudalis cerebelli (LCa) of the cerebellum; **(F-H)** lobus facialis (LVII) and lobus vagus (LX) of the medulla oblongata. Scale bar, 200 μm.

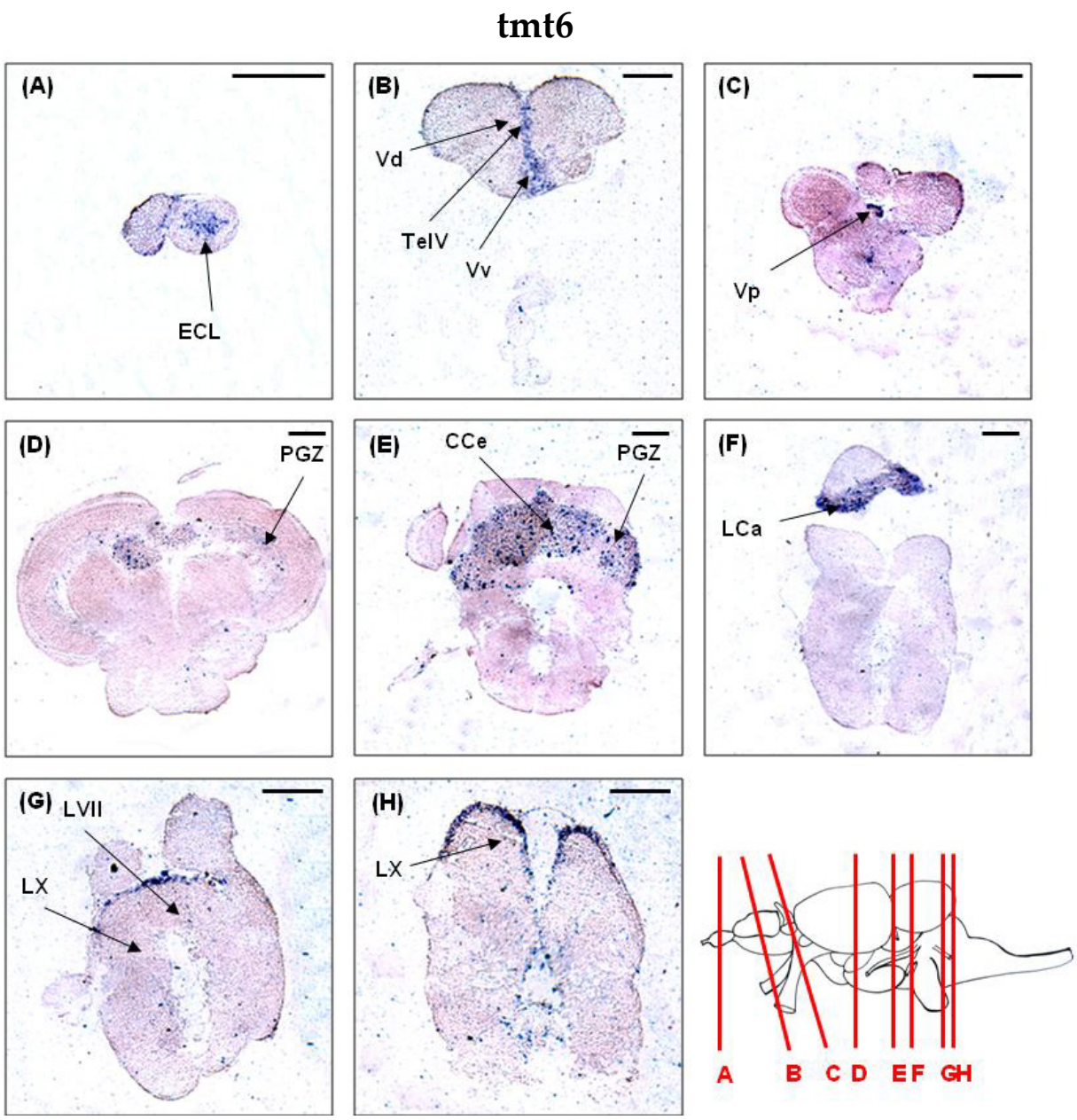


Figure 4.5 Expression of the *tmt6* gene in the adult zebrafish brain. *In situ* hybridisation detection of *tmt6* transcripts was performed on transverse rostro-caudal sections (A-G). *tmt6* expression was detected in: **(A)** external cellular layer (Vogt et al.) of the olfactory bulb; **(B)** the telencephalic ventricle (TelV), the dorsal (Vd) and ventral (Vv) nuclei of the ventral telencephalon; **(C)** postcommisural nucleus of the ventral telencephalon (Vp); **(D)** periventricular gray zone (PGZ) of the mesencephalon; **(E-F)** PGZ, corpus cerebellum (CCe), and lobus caudalis cerebelli (LCa) of the cerebellum; **(G-H)** lobus facialis (LVII) and lobus vagus (LX) of the medulla oblongata. Scale bar, 200 μm.

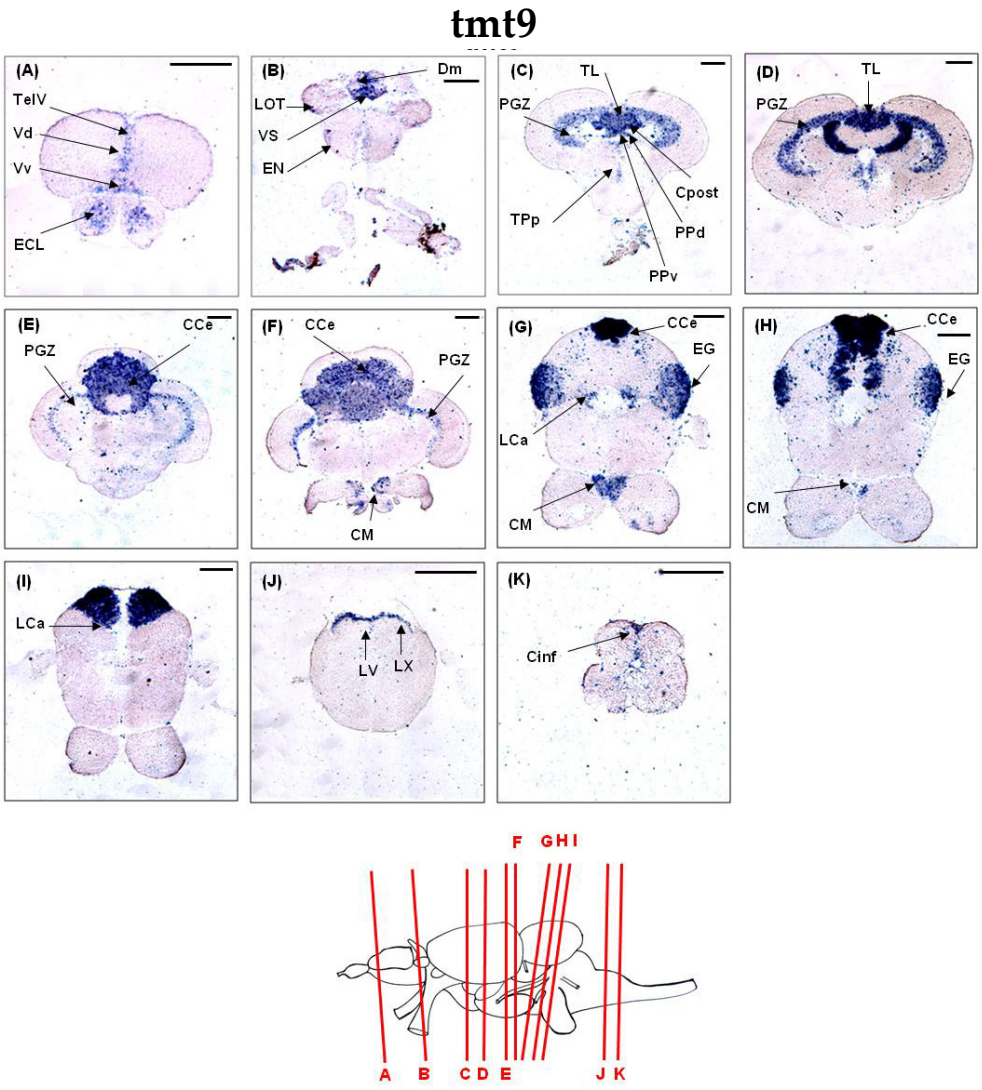


Figure 4.6 Expression of the *tmt9* gene in the adult zebrafish brain. RNA ISH detection of *tmt9* transcripts was performed on transverse rostro-caudal sections (A-K). *tmt9* expression was detected in: **(A-B)** the external cellular layer (Vogt et al.) of olfactory bulb, telencephalic ventricle (TelV) and dorsal (Vd) and ventral (Vv) nuclei, medial zone of D (Dm), supracommisural nucleus of V (Vs), entopeducular nucleus (EN), and the lateral olfactory tract (LOT) of ventral telencephalon; **(C)** periventricular gray zone (PGZ), torus longitudinalis (TL), dorsal (PPd) and ventral (PPv) portions of the periventricular prectectal nuclei, commissural posterior (Cpost), and the periventricular prectectal nucleus of the posterior tuberculum (Tpp); **(D-F)** PGZ, TL, corpus cerebellum (CCe), and corpus mamillare (CM) of the mesencephalon; **(G-I)** CCe, CM, eminentia granularis (EG), and lobus caudalis cerebelli (LCa) of the cerebellum; **(J-K)** lobus facialis (LVII), lobus vagus (LX) and comissura infima of Haller (Cinf) of the medulla oblongata. Scale bar, 200 μm.

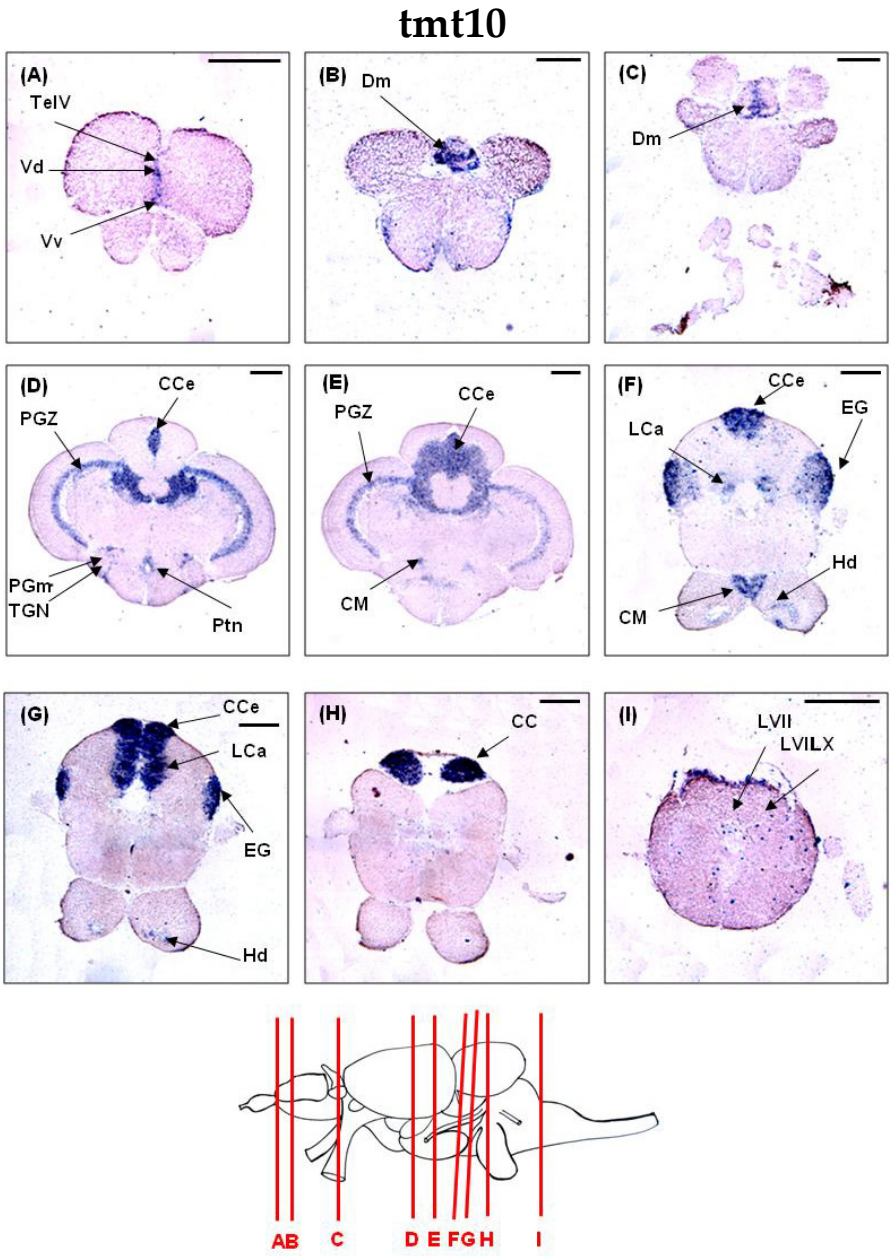


Figure 4.7 Expression of the *tmt10* gene in the adult zebrafish brain. RNA ISH detection of *tmt10* transcripts was performed on transverse rostro-caudal sections (A-I). *tmt10* expression was detected in: **(A-C)** the telencephalic ventricle (TelV), medial zone of D (Dm), dorsal (Vd) and ventral (Vv) nuclei in the ventral telencephalon; **(D-E)** periventricular gray zone (PGZ), corpus cerebellum (CCe), the medial periglomerular nuclei (PGm), the tertiary gustatory nucleus (TGN), and the posterior tuberal nuclei (PTN) of the mesencephalon; **(F-G)** CCe, corpus mamillare (Powell et al.), lobus caudalis cerebelli (LCa), eminentia granularis (EG), and dorsal zone of periventricular hypothalamus (Hd); **(H)** crista cerebellaris (CC); **(I)** lobus facialis (LVII) and lobus vagus (LX) of medulla oblongata. Scale bar, 200 μm.

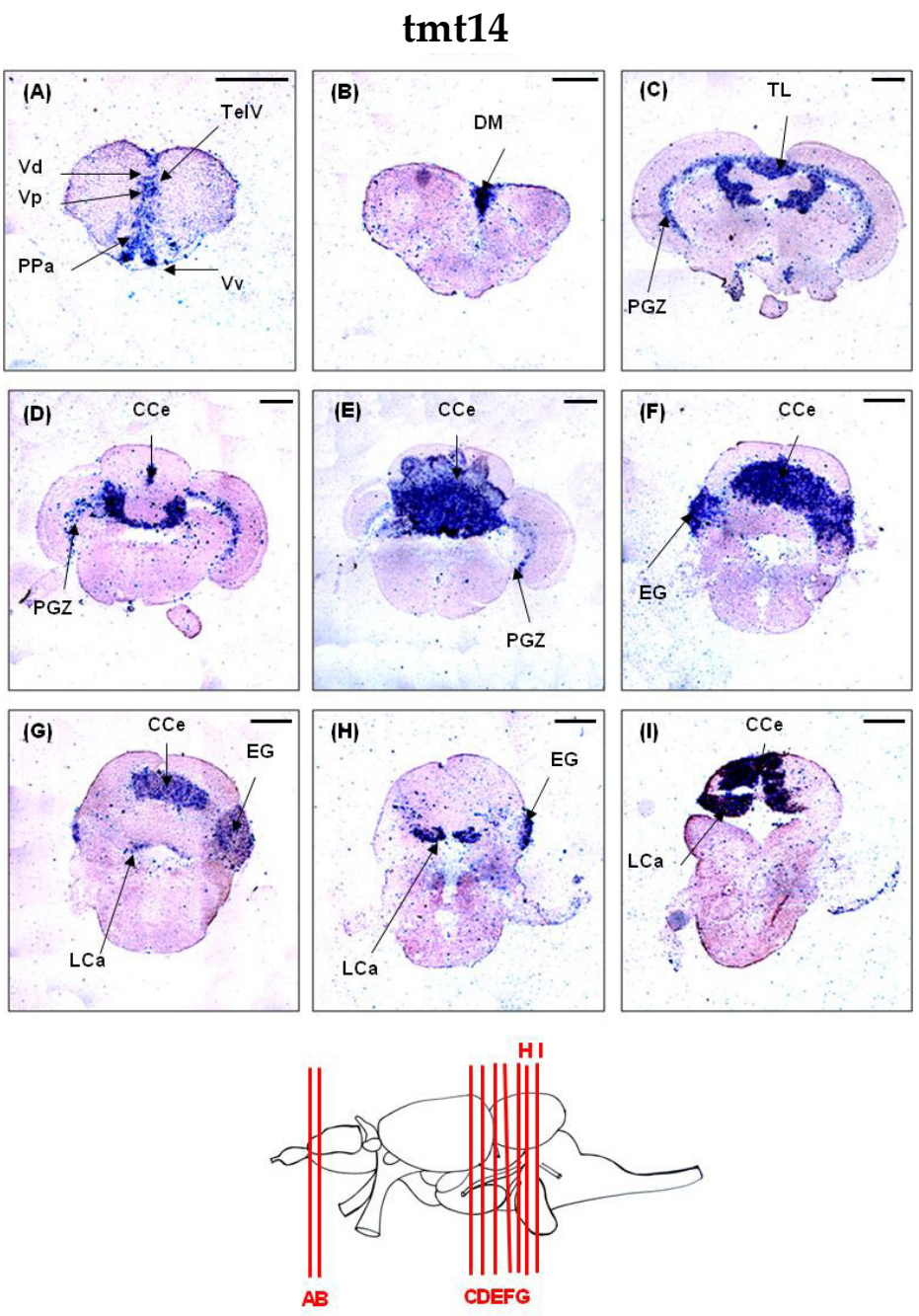


Figure 4.8 Expression of the *tmt14* gene in the adult zebrafish brain. *In situ* hybridisation detection of *tmt14* transcripts was performed on transverse rostro-caudal sections (A-I). *tmt14* expression was detected in: **(A,B)** the telencephalic ventricle (TelV), dorsal (Vd), ventral (Vv), and postcomissural (Vp) nuclei, and medial zone of D (Dm) in the telencephalon, and the parovocellular preoptic nucleus (Ppa) of the diencephalon; **(C-E)** torus longitudinalis (TL), periventricular gray zone (PGZ), and corpus cerebellum (CCe) of the mesencephalon and rhombocephalon; **(F-I)** CCe, eminentia granularis (EG), and lobus caudalis cerebelli (LCa) of the cerebellum. Scale bar, 200 μm.

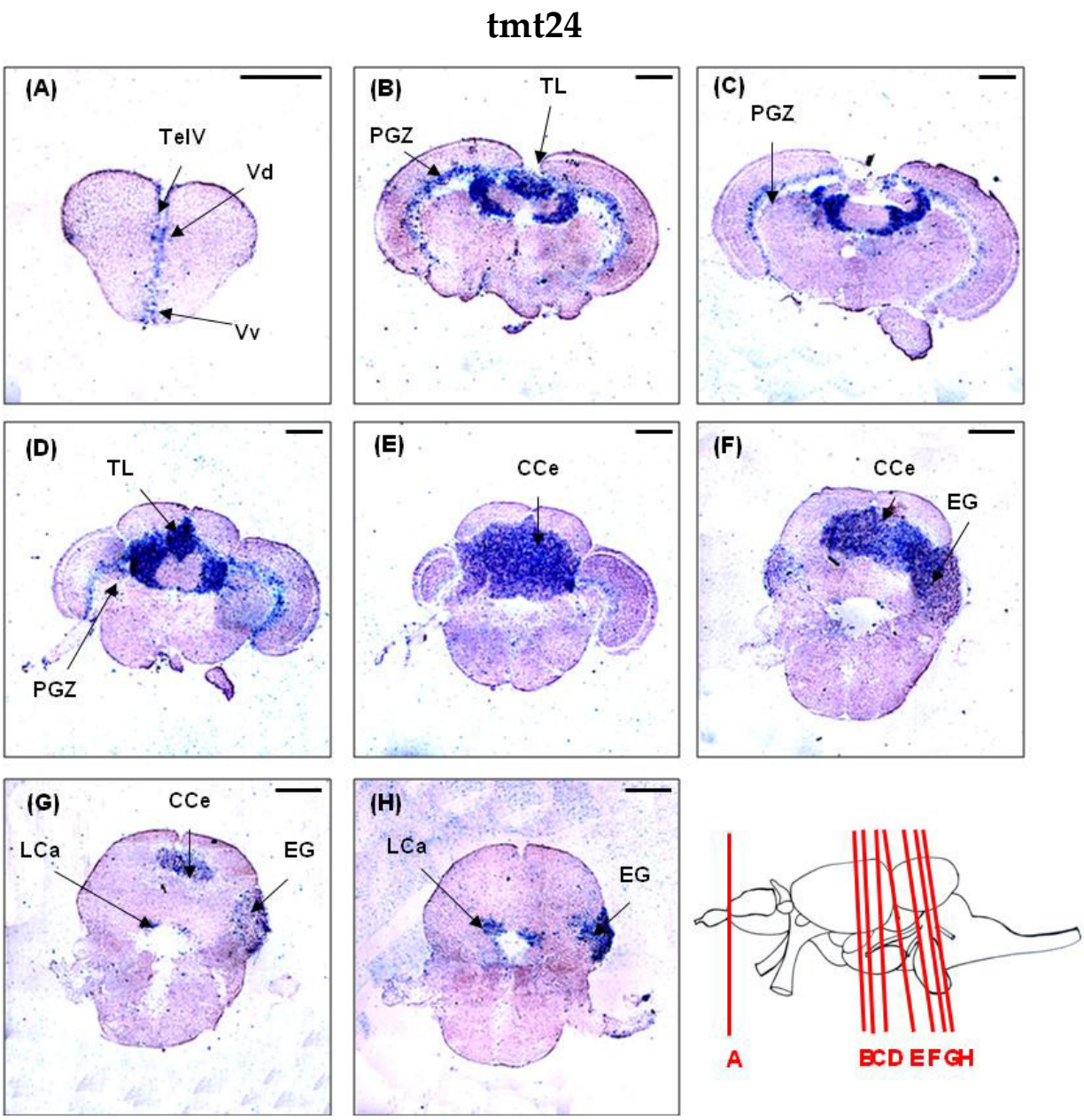


Figure 4.9 Expression of the *tmt24* gene in the adult zebrafish brain. *In situ* hybridisation detection of *tmt24* transcripts was performed on transverse rostro-caudal sections (A-H). *tmt24* expression was detected in: **(A)** the dorsal (Vd) and ventral (Vv) nuclei, and the telencephalic ventricle of telencephalon; **(B-D)** torus longitudinalis (TL), periventricular gray zone (PGZ), and corpus cerebellum (CCe) of the mesencephalon; (D) corpus cerebellum (CCe) of the rhombocephalon; **(F-H)** CCe, eminentia granularis (EG), and lobus caudalis cerebelli (LCa) of the cerebellum. Scale bar, 200 μm.

Table 4.1 Summary table of *tmt* opsin mRNA expression in the adult zebrafish brain.

	tmt2	tmt6	tmt9	tmt10	tmt14	tmt24
Telencephalon						
OB; Olfactory Bulbs						
ECL; external cellular layer of olfactory bulb		X	X			
TelV; Telencephalic ventricle	X	X	X	X	X	X
D; Dorsal telencephalic area						
Dm; medial zone of D			X	X	X	
LOT; lateral olfactory tract			X			
V; Ventral telencephalic area						
Vd; dorsal nucleus of V	X	X	X	X	X	X
Vv; ventral nucleus of V	X	X	X	X	X	X
Vp; postcomissural nucleus of V			X			
Vs; supracommisural nucleus of V			X			
EN; entopeducular nucleus			X			
Diencephalon						
Area preoptica						
Ppa; parovocellular preoptic nucleus, anterior part					X	
Thalamus	No expression observed					
Pretectum						
PPd; periventricular prectectal nucleus, dorsal part	X		X			
PPv; periventricular prectectal nucleus, ventral part	X		X			
Cpost; commissura posterior			X			
Posterior tuberculum						
Tpp; periventricular nucleus of the posterior tuberculum			X			
CM; corpus mamillare			X	X		
PGm; medial preglomerular nucleus				X		
Hypothalamus						
Hd; dorsal zone of periventricular hypothalamus				X		
Mesencephalon						
Superior and inferior colliculi						
PGZ; periventricular gray zone	X	X	X	X	X	X
TL; torus longitudinalis			X		X	X
Nmlf; nucleus of the medial longitudinal fascicle			X			
Tegmentum						
TGN; tertiary gustatory nucleus				X		
Rhombocephalon						
Cerebellum						
CCe_{mol}; corpus cerebellum	X	X	X	X	X	X
EG; eminentia granularis	X		X	X		X
LCa_{mol}; lobus caudalis cerebelli	X	X	X	X	X	X
Medulla oblongata						
LVII; lobus facialis	X	X		X		
LX; lobus vagus	X	X				
CC; crista cerebellaris				X		
Cinf; comissura infima of Haller			X			

Consistent with most RT-PCR data, the results of RNA ISH demonstrated that all *tmt* opsins are expressed in the adult brain of zebrafish. The only anomaly was the expression of *tmt24* not being detected at all by RT-PCR, suggesting variability in its amplification reaction. Overall, RNA ISH data revealed a much more complex expression pattern of *tmt* opsins in the adult brain of zebrafish compared to their developing brain in the embryos. Early expression of most *tmt* opsin genes (*tmt2*, *tmt6*, *tmt9*, and *tmt10*) seems to be confined to the medulla oblongata during the first five days of embryogenesis. Mapping their locations in the adult brain tissues may therefore offer clues on how they orchestrate different biological functions in the zebrafish (discussed below).

4.2.5 Quantification of *tmt* opsin gene expression by NanoString nCounter

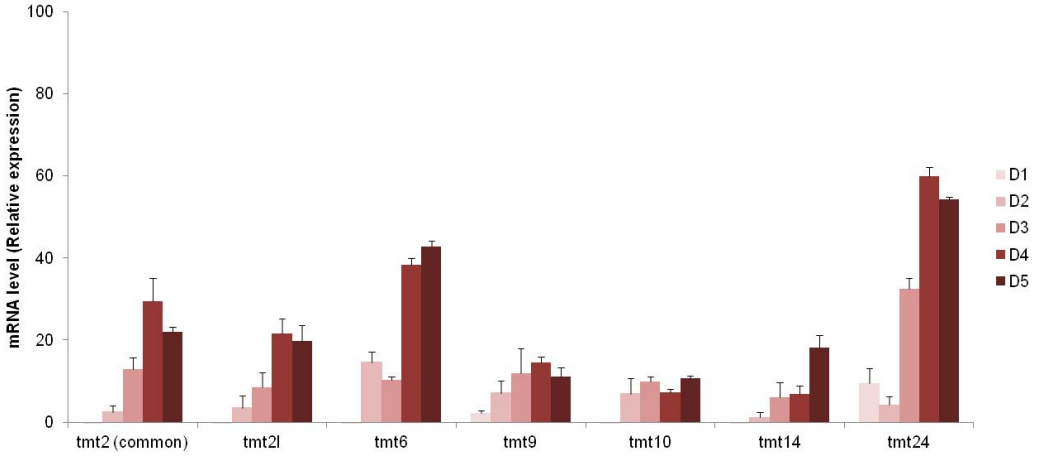
NanoString nCounter analysis system was used as a secondary validation assay to quantify the relative mRNA levels of *tmt* opsins in zebrafish day 1 to day 5 (D1-D5) embryos and adult tissues (brain, eye, testis, heart, gut, gill, and fin) (see Section 2.4.3 for details of methods and primers). One of the primers (*tmt2 common*) was specifically designed from the exon2 of *tmt2* to detect both its long and short isoforms, as *tmt2s* does not have exon4 and its exon3-intron3 is too restricted to encompass a divergent region for binding of the primer.

During the first five days of zebrafish embryogenesis, *tmt9* and *tmt24* were the earliest *tmt* transcripts detected from day 1 onwards (Fig. 4.10A). Although *tmt9* and *tmt24* transcripts were not observed in day 1 embryos from whole-mount ISH, this could be due to their low level of expression below the threshold for detection. All the other *tmt* opsins, *tmt2*, *tmt6*, *tmt10*, and *tmt14* were detected from day 2 onwards.

There was no significant differences observed between the expression of *tmt2 (common)* and *tmt2l*, suggesting that *tmt2s* might only be present at a very low level throughout the developmental period. The expression of *tmt2l*, *tmt6*, and *tmt24* had a 2- to 3-fold increase from day 3 to day 4, whilst that of *tmt14* had a 2-fold increase from day 4 to day 5. The transcripts of *tmt9* and *tmt10* remained at a relatively low, constant level from day 2 to day 5. In general, NanoString data and whole-mount ISH analyses showed a coherent expression pattern of *tmt* opsins during the early development of zebrafish embryos. However, not all of the *tmt* opsin transcripts were detected from the previous RT-PCR experiment (Section 4.2.1), which could be due to low sensitivity of the technique and the quality of cDNA samples.

Examination of *tmt* opsin expression in adult zebrafish tissues demonstrated that all six *tmt* opsins (*tmt2l*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, and *tmt24*) were present in the brain, testis, heart, and gill (Fig. 4.10B). In other tissues, different combinations of the *tmt* opsins were detected: eyes (*tmt2l*, *tmt6*, *tmt9*, *tmt10*, and *tmt24*); gut (*tmt6* and *tmt9*); fin (*tmt6*, *tmt9*, *tmt10*, *tmt14*, and *tmt24*). The expression level of *tmt2 (common)* was not significantly different to that of *tmt2l* in all the tissues tested, suggesting that the short isoform of *tmt2* was either absent or only weakly expressed. Most *tmt* opsins (apart from *tmt6* and *tmt10*) were found to be more abundant in the brain than in other tissues, which were also supported by the RT-PCR data (Section 4.2.1). Interestingly, *tmt6* transcript was shown to be the most highly expressed *tmt* opsin in the eye and testis of zebrafish, indicating that it may play an important role in the function of these tissues. All in all, NanoString analysis confirmed that the mRNAs of most *tmt* opsins are expressed in the tissues examined.

(A)



(B)

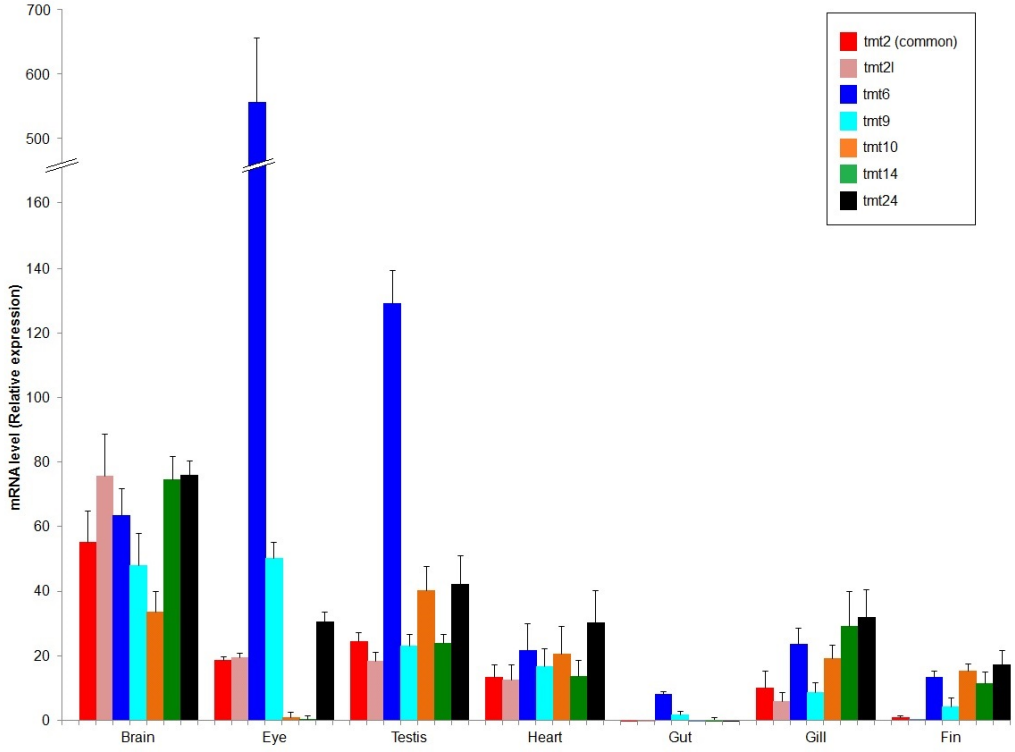


Figure 4.10 NanoString analysis of the expression of six *tmt* opsin genes (*tmt2_common*, *tmt2l*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, and *tmt24*) in **(A)** day 1- day 5 (D1-D5) zebrafish embryos and **(B)** adult zebrafish tissues including brain, eye, testis, heart, gut, gill, and fin. Data of *tmt2_common* represents the level of expression for both *tmt2l and tmt2s*, as the primer was designed to detect long and short isoform variants. The gene expression measurements of *tmt* opsins were normalized to the geometric mean of three endogenous controls *RPL13A*, *EF1A* and *ubiquitin B*. All NanoString data are presented as mean ± standard error of the mean (SEM).

4.2.6 Promoter analysis of zebrafish tmt opsins by bioinformatics

RT-PCR and RNA ISH data suggest an overlapping but distinct pattern of expression for each *tmt* opsins, therefore tissue-specific and temporal changes may be reflected by different promoter elements of the opsins. In an attempt to identify homologues of *tmt* opsins that may have extended sequences for examination of their promoter region, the *D. rerio* expressed sequence tag (EST) database was probed with full-length nucleotide sequences of the zebrafish *tmt* opsins using blastn algorithm at http://blast.ncbi.nlm.nih.gov/. However, no opsin-matched sequences were found. 5′ RACE PCR (see Section 2.4.4 for detailed method) was subsequently used to determine the transcription start site (TSS) of *tmt* opsins isolated from eye and brain tissues, and thus the location of their promoter. By sequencing (as described in Section 2.4.5) and comparison with the full-length sequences obtained previously, it was revealed that only fragments of *tmt2* and *tmt6* were successfully amplified. These two fragments were aligned with the genomic sequences containing the *tmt* opsin genes, and promoter analyses were performed 3 kb upstream of their coding regions. The schematic diagrams of the organization of their promoter regions are illustrated in Figure 4.11, whilst detailed annotations are shown in appendix A.2.

Examination of the 5′ promoter regions of *tmt2* and *tmt6* opsins identified a single TSS at 123 bp and 197 bp, respectively, upstream of the first nucleotide of the main translation initiation codon (mAUG). Proximal promoter region within about 200 bp upstream of the TSS are known to contain multiple transcription factor binding sites (*cis*-regulatory elements) that are critical for initiation of gene expression (Lodish, 2004, Griffiths, 2000). In both *tmt2* and *tmt6* genes, the proximal

promoter elements identified include TATA box, Downstream Promoter Element (DPE), GC box, and AP-1 (Fig. 4.11). A CpG island (a general characteristic of genes where promoters are likely to be located) and a B Recognition Element (BRE) were also identified in *tmt2* transcript (Fig. 4.11A), but were not found in *tmt6* transcript (Fig. 4.11B).

It is known that eukaryotic genes typically contain other long-range DNA regulatory elements (e.g. enhancer sequences) that can influence gene expression over the distances of 3 kb or more from the TSS (Noonan and McCallion, 2010, Hwang et al., 2013). Since these long-range regulatory elements are usually involved in mediating developmental or tissue-specific expression of the genes (Kleinjan and van Heyningen, 2005), distal promoter region encompassing up to 3 kb upstream of the TSS was examined in *tmt2* and *tmt6*. Additional binding sites for transcription factors that are associated with gene regulation in the eye, heart, liver, the lung and testes were identified (Fig. 4.11). Those involved in ocular gene expression include: central homeobox genes that control retinal development (e.g. paired box genes; pax-4 and pax-6) (Walther and Gruss, 1991, Wawersik and Maas, 2000); those that regulate the development of photoreceptors (e.g. rar-related orphan receptor beta (rorβ) (Jetten et al., 2001), cone-rod homeobox (crx) (Furukawa et al., 1997), retina-specific region-1/photoreceptor conserved element-1 (ret-1/pce-1), orthodentical homeobox/bovine AT-rich sequence-1 (otx/bat-1) (Kimura et al., 2000), retina-specific region-4 (ret-4) (Chen and Zack, 1996), glass-like binding motif (Moses and Rubin, 1991)); and those that mediate the differentiation of bipolar cells (e.g. *Caenorhabditis elegans* homeobox-10 (ceh-10)) (Dorval et al., 2006). As for regulatory elements that

are known to be specific for expression in other tissues, these include: homeobox gene nkx2.5, which is essential for heart formation and development (Balci and Akdemir, 2011, Guner-Ataman et al., 2013); hepatocyte nuclear factors (e.g. HNF-4), which are crucial for liver function but are also expressed in gut, kidney and pancreatic beta cells (Duncan et al., 1994); E26-like transcription factor 1 (elk-1), which contributes to signalling in the brain (Sgambato et al., 1998), as well as development of the liver and testis (Rao et al., 1989). Moreover, analysis also revealed the presence of enhancers that are important in the regulation of circadian-related genes, such as d-box and e-box (Munoz and Baler, 2003, Zhang et al., 2004), suggesting that *tmt* opsins may be directly linked to the circadian clocks in central or peripheral system at the transcriptional level. Taken together, these results are consistent with the wide spread expression of tmt opsins in zebrafish. Coupled with the putative presence of clock-related regulatory elements in the promoters, these data provide further evidence consistent with the hypothesis that tmt opsins may be important components of the peripheral clock system in the zebrafish.

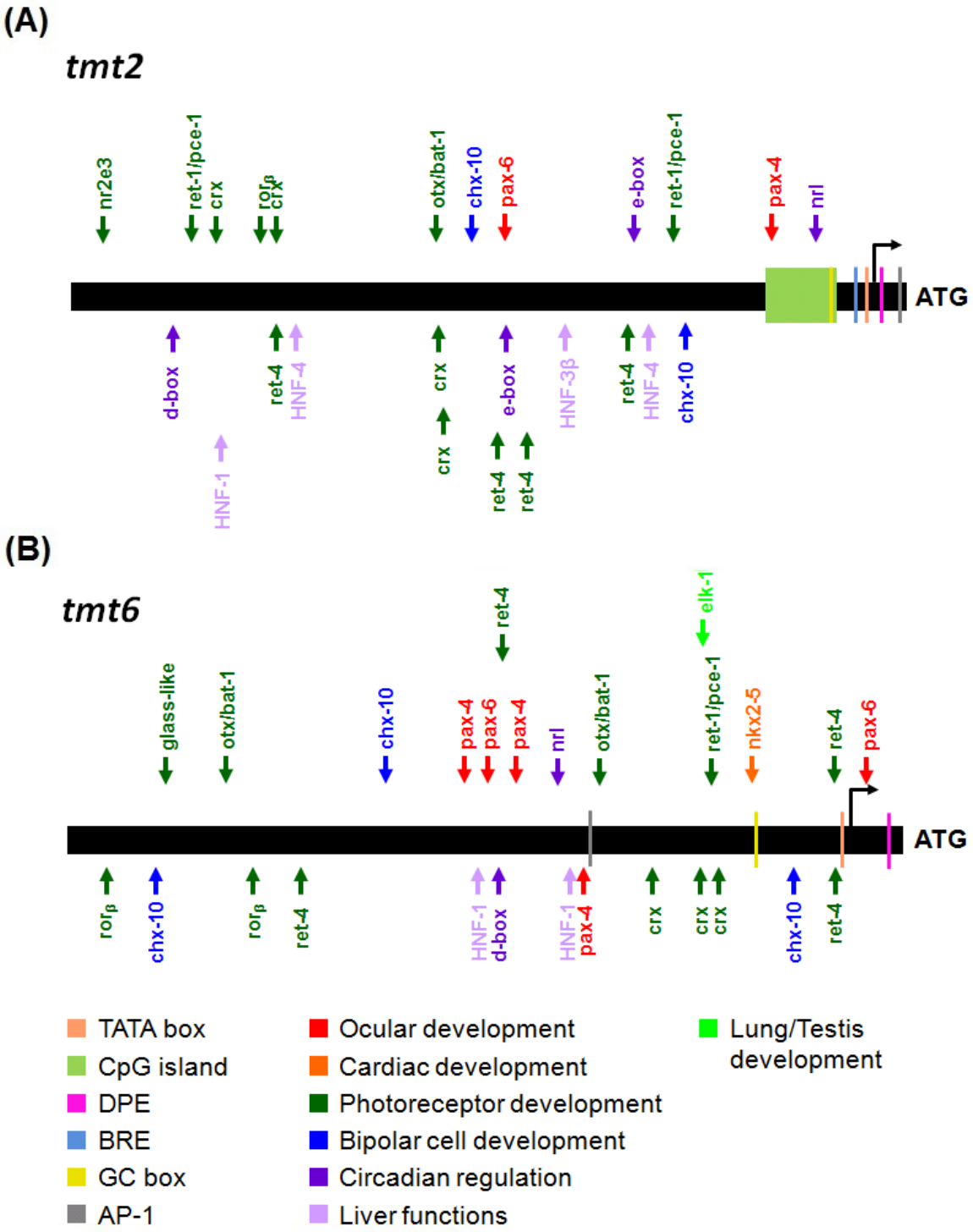

(A)
tmt2
nr2e3
ret-1/pce-1
crx
rorβ
crx
otx/bat-1
chx-10
pax-6
e-box
ret-1/pce-1
pax-4
nrl
ATG
d-box
HNF-1
ret-4
HNF-4
crx
crx
e-box
ret-4
ret-4
HNF-3β
ret-4
HNF-4
chx-10
(B)
tmt6
glass-like
otx/bat-1
chx-10
ret-4
pax-4
pax-6
pax-4
nrl
otx/bat-1
elk-1
ret-1/pce-1
nkx2-5
ret-4
pax-6
ATG
rorβ
chx-10
rorβ
ret-4
HNF-1
d-box
HNF-1
pax-4
crx
crx
crx
chx-10
ret-4
TATA box
CpG island
DPE
BRE
GC box
AP-1
Ocular development
Cardiac development
Photoreceptor development
Bipolar cell development
Circadian regulation
Liver functions
Lung/Testis development

Figure 4.11 Promoter analyses of 5′ flanking region (3 kb) of zebrafish *tmt2* and *tmt6*. Using the data obtained from 5′ RACE PCR and sequencing, the main transcription start sites were localized to -123 bp for **(A)** *tmt2* and -90 bp for **(B)** *tmt6*, upstream to the main translation initiation site (mAUG) which was denoted as position +1. Promoter analysis for each gene was performed within a region of 3 kb upstream to the main TSS. Firstly, CpG islands and CCCTC-binding factor (CTCF) insulator element were mapped, but only one CpG island was identified in *tmt2*. A putative TATA-box (marked by an arrow with the direction of transcription) is present in the promoter of both *tmt* opsins, at 16 bp (*tmt2*) and 20 bp (*tmt6*) immediately upstream of the main TSS. At the promoter, binding sites for transcription factors that are generally involved in basal gene regulation were also mapped. These include CAAT-box, GC-box, AP-1, and Oct-1. In addition, the promoter regions were analysed for tissue-specific regulatory sites, where putative motifs consistent with gene regulation in the development of various tissues were determined. These include the homeobox genes essential for retinogenesis (e.g. pax-4 and pax-6) (Grindley et al., 1995, Kozmik, 2005), factors regulating liver functions (e.g. HNF-1, HNF-3β, HNF-4) (Costa et al., 2003) and lung/testis development (e.g. elk-1) (Rao et al., 1989), genes involved in development and differentiation of photoreceptors (e.g. crx, ret-1/pce-1, otx/bat-1, rorβ, ret-4, glass-like, nr2e3, nrl) (Zhu and Craft, 2000, Swaroop et al., 2010, Akimoto, 2005) and bipolar cell (e.g. chx-10) (Dorval et al., 2006), as well as the sites which are important for circadian clock functions (e.g. d-box and e-box) (Zhang et al., 2004). All these binding sites are consistent with the broad expression of tmt opsins.

The sequences flanking the translation initiation codon (AUG) in the promoter region of eukaryotic mRNAs are known to be conserved across different organisms (Kozak, 1987, Lutcke et al., 1987). For vertebrates, the consensus sequence in this region, GCC$\begin{matrix}A\\G\end{matrix}$CC**AUG**G (Kozak, 1991a), has been shown to provide the strongest binding signal for ribosomal subunits and thus influence translation efficiency. Sequence analysis showed that *tmt6* has a slightly higher level of identity (60%) to the Kozak consensus sequence than *tmt2* (50%) (Fig. 4.11C), suggesting that perhaps translation from the mORF of *tmt6* is more efficient than that of the *tmt2* transcript. To measure the translation efficiency of *tmt2* and *tmt6* promoters, promoter-reporter fusion constructs can be generated with widely-used reporter genes such the green fluorescent protein (*gfp*) (Ali and Murrell, 2009) or the firefly luciferase (*luc*) (Vatine et al., 2009). The expression level of the fusion gene construct, which is driven by the activity of the specific promoter, can then be quantified in a heterologous expression system using a luminescence counter.

Table 4.2 Comparison of the 5′ flanking region of translation initiation codon in tmt2 and tmt6 with Kozak consensus sequence.

Start codon	Flanking sequence	Identity (%)*
Kozak consensus	-3 -1 A (GCC)GCC CC**AUG**G G	100
mAUG$_{tmt2}$	(CGG)AUUGCG**AUG**A	50
mAUG$_{tmt6}$	(GCG)GCAAAG**AUG**U	60

*Percentage homologies from positions -9 to +3 are indicated.

4.3 Discussion

Moutsaki and colleagues (2003) previously showed that zebrafish tmt opsin (*tmt2s*, accession number AF349947) was first discovered in peripheral tissues including heart, kidney, and in the PAC2 cell-line which displays photoentrainable circadian clock gene expression (Moutsaki et al., 2003). The work presented in this chapter has confirmed the expression of all six *tmt* opsins by RT-PCR (Section 4.2.1) and NanoString (Section 4.2.2) at different developmental stages of zebrafish embryos as well as in both peripheral and CNS tissues of adult zebrafish. Although there was the discrepancy between the RT-PCR results and those of *in situ* hybridisation, these qualitative data were verified by quantitative measure using NanoString analysis. It was found that all *tmt* opsin genes are expressed in a wide range of tissues, such as eye, brain, testis, heart, gut, gill and fin (Fig. 4.2A), which is consistent with the published results for *tmt2*. Additionally, their expression patterns were developmentally regulated (Fig. 4.2B). These findings, together with the structural analysis of zebrafish tmt opsins in Section 3.2, indicate a photosensory function that may be linked to circadian entrainment for this opsin subfamily.

Whole-mount ISH and NanoString data revealed that during the early development of zebrafish from day 1 to day 5, the expression of *tmt* opsins are temporally and spatially regulated. Some of the *tmt* opsin transcripts (*tmt9* and *tmt24*) were detected by NanoString from day 1 onwards (Fig. 4.10). Zebrafish embryos can detect light from as early as six to nine hours post fertilization (hpf) (Tamai et al., 2004), thus providing evidence that supports the potential light-detecting ability of *tmt* opsins. The observation of all *tmt* opsins being initially

expressed in the notochord and neural tubes suggests that they may play an important role in the development of the CNS from an early developmental stage (Fig.4.2). Labelling of riboprobes was also observed in the neurocranial trabecular region for *tmt2*, in the medulla oblongata for *tmt2*, *tmt6*, *tmt9*, and *tmt10*, and in the gut and liver for all *tmt* opsins studied. The neurocranial trabecular region is part of the neurocranium that forms the olfactory system, whilst the medulla oblongata (also known as hindbrain) is known for controlling coordination, movement, and balance (Moens and Prince, 2002). Interestingly, rhythmic expression of clock genes has also been reported in the hindbrain, gut and liver of zebrafish (Lopez-Olmeda et al., 2010, Whitmore et al., 1998), indicating the possibility of overlapping functions between *tmt* opsins and clock genes in these regions.

In adult retina, RNA ISH showed that *tmt* opsins are expressed in many, but not all, retinal cell types. Their expression was most prominent in bipolar cells, where all six *tmt* transcripts were detected (Fig. 4.3). Detailed observation indicated that each of the opsin genes is expressed at varying levels in the bipolar cells. Bipolar cells are the last type of neurons to differentiate in the retina of zebrafish, forming only after 2 days post fertilization (dpf) (Schmitt and Dowling, 1999) and mature by 7 dpf (Biehlmaier et al., 2003). However, the transcripts of *tmt* opsins were not detected in the eye of day 2-5 zebrafish embryos from whole-mount ISH. It is known that retinal development can be dictated by light induction, which has been evident in the mouse eye (Rao et al., 2013). Since *tmt* opsins do not seem to be present in the eye of early developing zebrafish embryos, other opsins (e.g. opn4) are likely be responsible for controlling these developmental processes of the eye. It has been found that indeed

light detection by opn4-expressing RGCs play an important role in vascular patterning of the eye (Rao et al., 2013). Although the expression of multiple *tmt* opsins may not be associated with retinogenesis, they may still have an influence in the differentiation of different bipolar cells at later stages. There are 17 morphologically distinct types of bipolar cell currently described, which have been classified into six different types (S1-S6) based upon the location of their axon terminals in the inner plexiform layer (IPL) (Connaughton et al., 2004). Physiologically, the subtypes can be divided further into three groups depending on their change in membrane potential when stimulated by neurotransmitter glutamate: (1) 'group a' monostratified cells (ON-types, B_{on}S1-3) have single axon terminal ramifying in sublamina a of the IPL, and they respond to light by depolarization; (2) 'group b' monostratified cells (OFF-type, B_{off}S4-6) have single axon terminal confined to sublamina b, and they respond to light by hyperpolarization; or (3) 'group a/b' multistratefied (ON- or OFF-types) cells with multiple axon terminals in both sublamina a and b. The expression of each *tmt* opsin might be specific to a bipolar cell type, or equally likely multiple *tmt* opsins can be co-expressed in some of the cell types, providing coordinated control of light sensitivity between different cells. Weak expression of *tmt*2, *tmt9*, *tmt10*, and *tmt14* was detected in the retinal ganglion cells (RGCs), whilst intense expression of *tmt6* was observed in a subtype of amacrine cells (Fig. 4.3A). Both of these regions have been linked with intrinsic photosensitivity that may contribute to the regulation of circadian rhythms in the teleost eye (Zele et al., 2011, Davies et al., 2011). Although melanopsin has already been implicated as the candidate photopigment that may mediate these responses in

the retina (Schmidt et al., 2011, Davies et al., 2011), the presence of tmt opsin may also play a part. To distinguish the exact subtypes of bipolar cells, RGCs and amacrine cells that express tmt opsins, immunocytochemistry using antibodies to target them, removal of their functions by siRNA/morpholinos or generation of knock-out mice would be useful.

The translucence of extra-cranial and cranial tissues in the zebrafish easily allows light to penetrate through to different regions of the brain. Significantly, light can exert direct effect on cranial functions, by up-regulating the expression of various genes or controlling the release of neurotransmitters (Weger et al., 2011, Jones, 2007). ISH of *tmt* opsins in adult zebrafish brain demonstrated that each of the opsin genes shows a widespread, yet complex pattern of expression (Fig. 4.4-4.9). Some of their localisations correlate with the reported findings of brain regions that are photoreceptive. Within the preoptic area, transcripts of *tmt* opsins were detected in the ventral havenula (*tmt6*, *tmt14*), anterior part of the parvocellular preoptic nucleus (*tmt6*, *tmt14*), dorsal and ventral portions of the periventricular pretectal nucleus (*tmt2*, *tmt9*). Expression of *tmt* opsin were also localised to various regions at the base of the hindbrain, e.g. corpus cerebella, lobus caudalis cerebella, and medulla oblongata. Both the preoptic and the basal ganglia regions (including cerebellum and brain stem) have been implicated in the role of deep brain photoreception and circadian entrainment (Foster et al., 1994, Vigh et al., 1983). Importantly, *tmt* opsins were also localised to other regions of the brains that are not known to possess a role in circadian photoreception, indicating that they may have other biological functions. These areas include: the external cellular layer of the olfactory bulb and lateral

olfactory tract, which are involved in detection of odour and feeding behaviours (Tabor et al., 2004, Hansen and Zeiske, 1998); the ventral havenula and parvocellular preoptic nucleus in the preoptic tectum, which are concerned with movement and shape (Douglas and Djamgoz, 1990); the vagal lobe, cerebellar granular eminence, and corpus cerebella, which are part of the mechanosensory circuitry for balancing and movements (Wagner, 2001); facial and vagal lobes in the medullary gustatory column which, together with the tertiary gustatory nucleus in the midbrain, form part of the gustatory system (Kanwal and Finger, 1997). Therefore, it seems that tmt opsins may be crucial in diverse array of biological functions.

As part of the investigation, regulatory mechanisms that mediate ontogenic and spatial differentiation of *tmt* opsins expression were predicted by analysing the promoter region of each tmt gene. Amongst the predictions, putative motifs within the proximal region upstream of the transcription start sites were identified (e.g. TATA box, DPE, GC box, and AP-1 (Fig. 4.11)) for *tmt2* and *tmt6*. As these are important for the recruitment of basal transcription factors, it would appear that this region may comprise the core promoter. Nuclear receptor binding sites of rorβ, crx, ret-1/pce-1, ret-4, ceh-10, nkx2.5, HNF-4 were also identified (Fig. 4.11), which are predominately known to control tissue-specific expression and are in agreement with the differential expression pattern observed from RT-PCR, Nanostring analysis and RNA ISH (Section 4.2.1-4.2.5).

In summary, the studies in this chapter demonstrated that expression of *tmt* opsins is developmentally and spatially regulated. Analysis of the promoters in representative *tmt* opsins (*tmt2* and *tmt6*) has provided valuable insights into the

mechanisms of differential gene regulation that underlie the broad expression patterns observed. However, in order to improve the understanding of these complex processes of gene regulation, further functional studies are required, such as the quantitative measurement of promoter activities, for example, by using reporter deletion assays whereby different promoter-luciferase fusion constructs are generated and their activity responses quantified (Horstmann et al., 2004, Xu et al., 2013).

CHAPTER 5

Spectral Tuning and Retinoid Usage of tmt Photopigments

CHAPTER 5: Spectral Tuning and Retinoid Usage of tmt Photopigments

5.1 Introduction

Sequence comparison of the six zebrafish tmt opsins with other vertebrate photopigments revealed that tmt opsins have the necessary features to act as photosensory pigments (see Chapter 3). All functional vertebrate opsin-based pigments are known to possess a lysine residue at site 296 in the seventh transmembrane segment, where the retinal chromophore binds via a Schiff base covalent linkage (Nathans and Hogness, 1983, Hargrave et al., 1984). This conserved Lys296 has been identified in the six tmt opsins isolated from zebrafish (Fig. 3.2) and is indicative of potential light absorbing capability of the opsin pigments since photosensitive retinal can attach to this site. To obtain an insight into the possible light-sensory function of this group of opsins, UV-Vis spectrophotometric analysis was performed with tmt pigments that were regenerated *in vitro* with 11-*cis* and all-*trans* retinoids. These two types of retinoids were tested because most vertebrate opsins have been shown to preferentially couple with 11-*cis* retinal (a naturally occurring ligand) in the dark (Shichida and Matsuyama, 2009), whilst bistable pigments (e.g. opn5, and peropsin) can also form a stable association with all-trans-retinal (Koyanagi et al., 2002, Tarttelin et al., 2003).

A central question in the photoactivation of tmt pigments is what are the molecular mechanisms that control the absorption of specific wavelength of light? λ_{max} defines the wavelength at which maximum absorption of light occurs for a specific compound. The unbound form of retinal chromophore in solution is known

to display a λ_{max} in the UV region between 370-400 nm depending on the isomer variant (Fisher and Weiss, 1974, Stoeckenius et al., 1979). However, when bound to an opsin, protonation of the retinal Schiff base linkage usually causes a spectral shift in the λ_{max} of the photopigment towards ~440 nm in the visible region (Pitt et al., 1955). The only photopigments which have been shown to possess an unprotonated Schiff base in a dark state are the UV-sensitive (UVS) SWS1 opsins (Kusnetzow et al., 2004). In these UVS pigments, the presence of a non-polar amino acid (Ile, Cys, Met, Ala, or Phe) at position 86, which is close to the Schiff base linkage have been implicated for preventing its protonation (Hunt et al., 2004). Other residues surrounding the retinal binding pocket, Ser or Cys at site 90 and Thr or Gln at site 93, are also known to have a synergistic effect in interfering with protonation of the Schiff base (Odeen and Hastad, 2003, Hunt et al., 2004). This could be due to polarity and chemical structure of the amino acids at site 86, 90 and 93, which repel water away from the Schiff base, thereby displacing their positive charges (Yokoyama et al., 2000). Interestingly, tmt opsins do not possess these amino acids that seem to be attributable to the unprotonated Schiff base. One of the aims in this study is therefore to provide an understanding of the spectral tuning in a representative tmt opsin (tmt10), making use of site-directed mutagenesis to modify known spectral tuning sites and perform spectrophotometric analysis on the mutants generated.

Numerous experimental studies have been carried out to elucidate the mechanism of spectral tuning in opsin-based photopigments in vertebrates, many of which were determined by mutating visual pigments from an array of different animal groups (Chinen et al., 2005, Rajamani et al., 2011, Davies et al., 2012c). Main

factors which influence spectral tuning of photopigments include (1) changes in charge or size of specific amino acid residues within the binding pocket (Rajamani et al., 2011); (2) the nature and positioning of the counterion, which acts to stabilize the positive charge of the protonated Schiff base (Terakita et al., 2004, Kakitani et al., 1985); and (3) conformational changes in the chromophore upon light stimulation (Sundaralingam and Beddell, 1972). By mutating individual amino acid at sites that have been shown to be important for spectral tuning, findings from this study help to decipher the photoactivation mechanism of tmt opsins.

As mentioned before, the biochemical function of the counterion in a photopigment molecule is to counteract the positive charge on Schiff base, stabilising the tertiary structure of the inactive dark state protein as well as contributing to the tuning of wavelength absorbance. A net neutral charge in the retinal binding pocket is crucial as the presence of any charged or polar moiety near the chromophore may perturb the energy transition process during light activation of photopigment. This research is primarily based on visual pigments, where the visible range is an important baseline from which other wavelengths are spectral tuned to. It may be that for non-visual opsins, e.g. opn5, parapinopsin and RGR, UV sensitivity plays a more central role (Yamashita et al., 2010, Wada et al., 2012, Hao and Fong, 1996). In all classical visual pigments, the counterion has been identified as a negatively charged glutamate residue located at a conserved site 113 in the third transmembrane segment (Sakmar et al., 1989). In many non-visual opsins and invertebrate opsins, however, their site 113 is usually occupied by an uncharged residue, either tyrosine or phenylalanine (Terakita, 2005). As for photoisomerase

opsins, e.g. the retinochrome (Terakita et al., 2000), either a methionine or histidine residue is present at this position. In cases where Glu113 is not present in the opsins, it has been suggested that the counterion is displaced to another glutamate residue at position 181, which is highly conserved across most opsins (with the exception of histidine in long wave sensitive (LWS) pigments and serine in VA opsin) (Terakita et al., 2004, Wang et al., 1993, Davies, unpublished data). The presence of a positively charged His181 together another positively charged Lys184 in LWS pigments form a Cl^- binding site, which leads to absorption of longer wavelengths by the bound chromophore, thereby causing a red shift in the λ_{max} of these pigments (Kleinschmidt and Harosi, 1992, Wang et al., 1993). For other opsins, evidence supporting Glu181 as their putative counterion were derived from mutagenesis studies in bovine RH1 (Terakita et al., 2000), which demonstrated that substituting the uncharged amino acid at site 113 do not alter its absorption spectra. However, mutating Glu181 would cause a spectral shift as the Schiff base becomes unprotonated. Hence, it seems likely that Glu181 is capable in acting as a counterion. Table 5.1 below summarises the amino acid residues identified at positions 113 and 181 across three different groups of opsins.

Table 5.1 *Amino acid at positions 113 and 181 across different opsins groups*

	Site 113	**Site 181**
Vertebrate visual opsins	Glu	Glu, except for His in LWS opsin
Vertebrate non-visual opsins and Invertebrate opsins	Tyrosine or Phe (except for Glu in pinopsin and exo-rod opsin, and Asp in opn3)	Glu (except for Ser in VA)
Non-GPCR photoisomerase opsins	Met or His	Glu

This study describes a molecular functional approach for the spectral characterisation of the tmt opsin proteins in the zebrafish. The aim was also to elucidate the fundamental mechanism regulating the spectral tuning of zebrafish tmt pigments, using combined site-directed mutagenesis studies and UV-Vis spectrophotometry. It was anticipated that results from the mutagenesis experiments would lead to the identification of specific amino acid-chromophore interactions that mediate any observed spectral shifts.

5.2 Results

5.2.1 Regeneration and spectral sensitivity of the zebrafish tmt opsins

Absorbance studies were performed with the seven zebrafish tmt pigments (tmt2s, tmt2l, tmt6, tmt9, tmt10, tmt14, and tmt24) to characterise their spectral sensitivities. Full-length coding sequences for each tmt opsin were expressed *in vitro* using HEK293T cells (refer to Section 2.5.2 for detailed methodology). The opsin proteins harvested were reconstituted with excess 11-*cis* or all-*trans* retinals. Photopigments of tmt were then purified by immunoaffinity chromatography using a CNBr-activated Sepharose column coupled with anti-RH1 1D4 antibody before subjected to spectrophotometric analysis (see Section 2.5.3.1). Dark spectra for each of the reconstituted pigments were determined from a broad range of wavelengths between 200–700 nm at room temperature. Following this, the pigment was either bleached by exposure to broad spectrum incandescent white light (for VS pigments) or denatured by acid treatment (for UVS pigments) before spectral measurements were repeated. Detection of a λ_{max} at ~380 nm in the light-bleached pigment reflects deprotonation of its Schiff base linkage as the bound 11-*cis* retinal is converted into

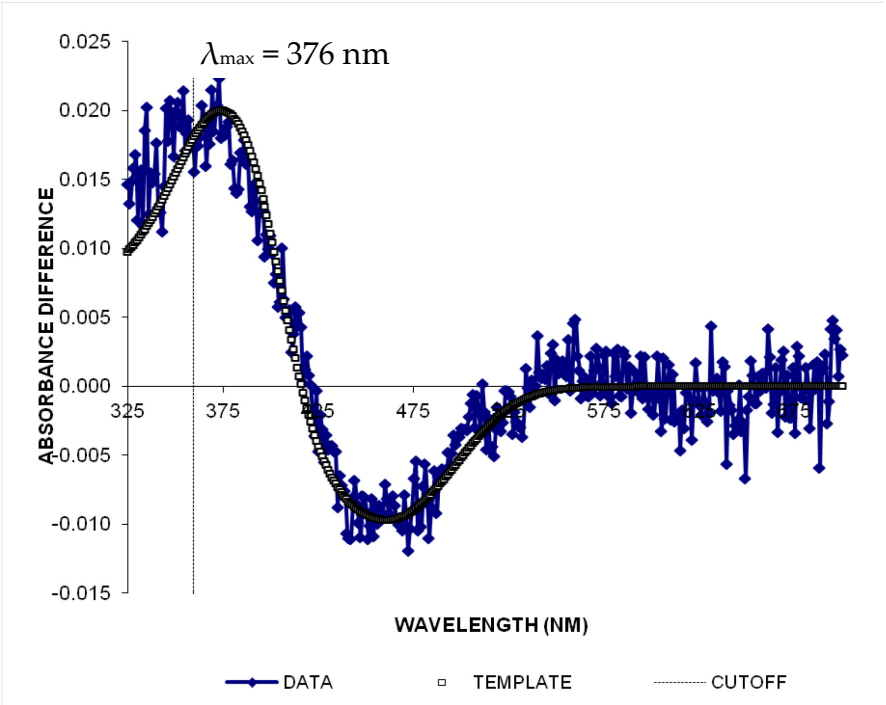
all-*trans* isomer (Farrens and Khorana, 1995, Nathans, 1990), indicating that the pigment is indeed functional. On the other hand, detection of spectral shift with a λ_{max} at 440 nm from the acid-treated spectra is indicative of protonated Schiff bases in the 11-*cis* retinals that are released from the denatured proteins (Sakamoto and Khorana, 1995, Kito et al., 1968). Difference spectra were then generated from the dark and bleached or acid-treated spectra, in order to account for the basal absorbance of protein and its effect on light scattering. These factors must be taken into consideration because they cause less light to reach the detector, leading to a falsely high absorbance value measured with elevated baseline. The wavelength of the peak absorbance, λ_{max}, for each regenerated pigment was determined by fitting a modified visual rhodopsin pigment template (Govardovskii et al., 2000, Parry et al., 2004, Davies et al., 2012b) to the difference spectra.

All the zebrafish tmt opsin pigments were successfully expressed and they appeared to give UVS dark spectra that shifted with acid treatment (data not shown). However, most of the data were rather noisy and the precise spectral peaks were difficult to calculate. The results from tmt10 alone gave substantial peaks above background, with λ_{max} at 376 nm when reconstituted with 11-*cis* retinal and at 368 nm with all-*trans* retinal (as shown in Fig. 5.1). Interestingly, a recent paper showed that homologue of tmt opsin from tmt class III in pufferfish (*Takifugu rubripes*) has a λ_{max} at ~460 nm in the dark when spectral measurements were taken at 4°C (Koyanagi et al., 2013). This value is significantly red-shifted compared to the ones measured for zebrafish tmt opsins, suggesting that there might be species-specific differences in the spectral tuning domains of the tmt pigments. The spectral differences observed

could also be due to variation in the experimental conditions, as temperature is known to have an effect on spectral absorbance of pigments (Ala-Laurila et al., 2003).

Taken together, these data confirmed the formation of stable UVS pigments, which bind 11-*cis* and all-*trans* chromophores, indicating bistability. Although further evidence on chromophore selectivity and photoreversion of the retinals is required to confirm the bistable nature of zebrafish tmt pigments, many other bistable pigments (e.g. *Branchiostoma belcheri* Amphiop1, *Callorhinchus milii* opn4, *D. rerio* opn4-m1 and opn4-m3) have been shown to form stable interactions with 11-*cis* and all-*trans* retinals independently (Koyanagi et al., 2002, Davies et al., 2011, Davies et al., 2012a). To investigate if zebrafish tmt opsins act as monostable or bistable photopigments, a series of functional experiments including whole-cell patch clamp recording, Ca^{2+} imaging, and luciferase-based bioluminescent reporter assays was performed with 9-*cis* as well as all-*trans* retinals in Chapter 6. 9-*cis* retinal is an artificial analogue of the native 11-*cis* retinal, and it was used as substitute for the native retinoid in all these functional experiments due to the large quantity required. Studies have shown that stability and activation of the pigments formed with 9-*cis* retinal closely resemble those with 11-*cis* retinal (Pepperberg et al., 1978, Van Hooser et al., 2000), thus the usage of either isomers is interchangeable.

(A)



(B)

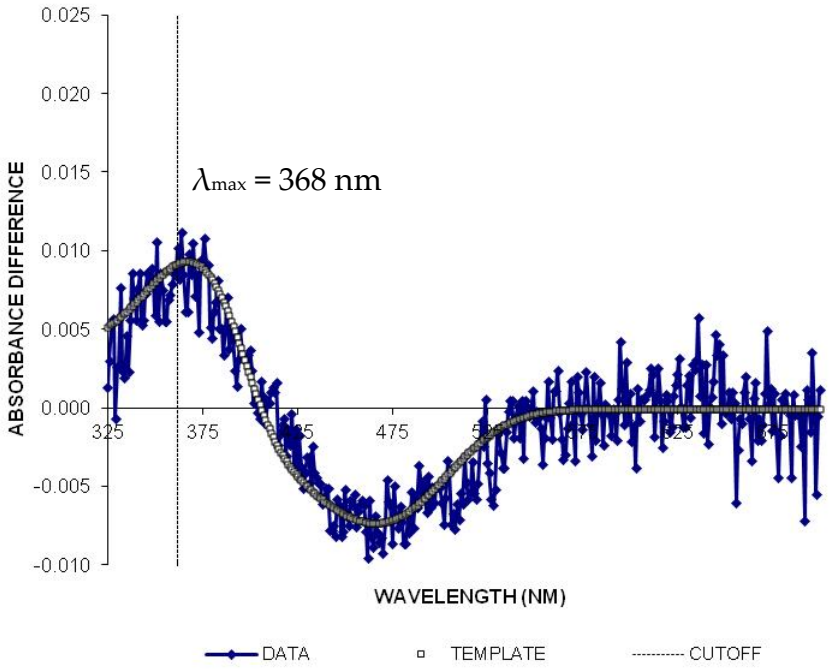


Figure 5.1 UV-Vis spectrophotometric analyses of zebrafish tmt10 reconstituted with **(A)** 11-*cis* and **(B)** all-*trans* retinals. The difference spectra were generated by subtracting the acid-denatured spectra from the dark absorbance spectra, which were then fitted against a modified visual rhodopsin template (Govardovskii et al., 2000, Parry et al., 2004) for determination of the peak absorbance (λ_{max}) values.

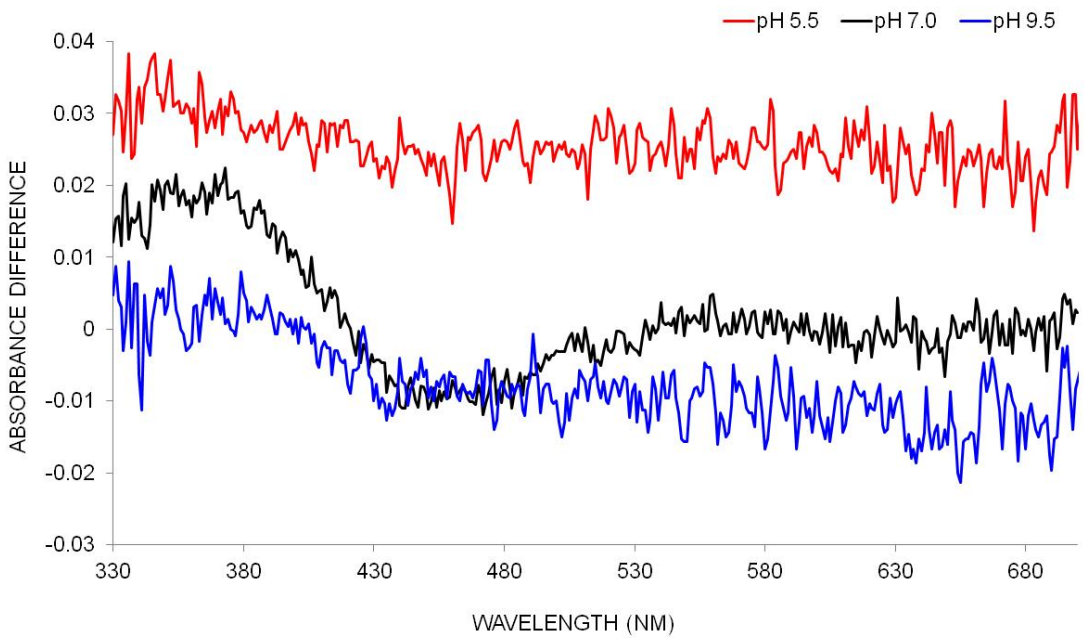
5.2.2 Protonation state of the retinylidene Schiff base in tmt10

To determine the protonation state of tmt10 pigments, pH-dependency of their absorption maximum was examined. The photopigments were regenerated with 11-*cis* retinal in buffer solutions under three different conditions for comparison: pH 5.5, pH 7.0 and pH 9.5 (see Section 2.5.3.3 for detail of the methodology). These pH values were selected based on their maximum effects in the spectral shift of other photopigments, e.g. peropsin and bovine RH1 (Terakita et al., 2004). It was predicted that the higher concentration of H^+ ion at acidic pH would induce protonation of the retinylidene Schiff base, whilst the alkaline pH would deprotonate this linkage in the pigment molecule.

Fig. 5.3B summarizes the λ_{max} of the absorbencies measured. Previous UV-Vis spectrophotometric analysis showed that at pH 7.0, tmt10 pigment reconstituted with 11-*cis* retinal had a strong absorbance peak at 376 nm. When tested at pH 5.5, although light stimulation of the tmt10 pigment caused the formation of a photoproduct with a small maximum at UV region (red line in Fig. 5.2A), acid treatment of the measured sample did not remove or shift this peak. Therefore, it seems that regeneration of the pigment has failed at pH 5.5. This suggests that increase of H^+ concentration destabilises the protein structure, possibly altering the charge density of its Schiff base linkage, which leads to displacement of the retinal from the binding pocket. In contrast, tmt10 tested at pH 9.5 (blue line in Fig. 5.2A) showed a difference spectrum (λ_{max}= ~370 nm) resembling that of the tmt10 measured at pH 7.0 (black line in Fig. 5.2A), suggesting that tmt10 is still a UVS opsin.

Assuming that the tmt10 pigment is unprotonated at ground state, decreasing pH of the buffers to 5.5 may promote protonation of the opsin, and thus shift its λ_{max} to the violet region, as seen in a goldfish and a mouse UVS mutant pigment (Yokoyama and Shi, 2000, Fasick et al., 2002, Hunt et al., 2004). If the pigment is already protonated, increasing the pH to 9.5 may result in unprotonation of the Schiff base and perhaps a shift further to even shorter wavelengths. The lack of H^+ ions at pH 9.5 resulted in no obvious change in spectral tuning of the tmt10 pigment, which suggests that it may be unprotonated in the first place. It is also possible that these tmt10 photopigments may be constantly protonated irrespective of the different pH conditions tested. Based on the limited data from this experiment, the protonation state of the tmt10 pigment cannot be confirmed.

(A)



(B)

	λ_{max} of the spectral absorbance measured
Zebrafish tmt10	**With 11-*cis* retinal**
pH 5.5	failed
pH 7.0	376 nm
pH 9.5	~370 nm

Figure 5.2 pH-dependent absorption spectral measurements in the photoproducts of zebrafish tmt10. **(A)** The graphs show difference absorption spectra of the wild-type tmt10 after light irradiation with wavelengths between 200–700 nm at pH 5.5 (red line), 7.0 (black line) and 9.5 (blue line). **(B)** A summary of the λ_{max} values deduced from the spectra of tmt10 pigments under different pH conditions.

5.2.3 Examination of putative spectral tuning sites by mutagenesis studies

UV-Vis spectrophotometry has demonstrated that zebrafish tmt pigments generally absorb in the UV region between 365-380 nm wavelengths. To examine the molecular bases for this UV sensitivity, amino acid sites that are known to have spectral tuning effects in UVS pigments were investigated by site-directed mutagenesis. Zebrafish tmt10 was selected as a representative tmt pigment from the group for this study, since its absorbance spectra gave the most well-defined UV peak. Specific amino acid locations that were tested include (1) site 113 and 181, which have been implicated for their role as counterions (Terakita et al., 2004); and (2) site 87, as cysteine is found to be conserved at this position in all zebrafish tmt and SWS1 pigments. The Cys87 is also located in close proximity to other sites that have been shown to be important for UV absorbance in SWS1 pigments, including site 86 (for non-avian SWS1 pigments), 90 and 93 (for avian SWS1 pigments) (Wilkie et al., 2000, Hunt et al., 2004).

Four different tmt10 constructs with single or double mutations were generated as described in Section 2.5.3.4 and analysed by spectrophotometry (data are summarised in Table 5.3G). These consisted of (1) Tyr113Glu, whereby the uncharged tyrosine was substituted with a negatively charged glutamate that might act as another counterion to the Schiff base in addition to Glu181; (2) Glu181Gln, removal of the putative counterion at this site by substituting the glutamate residue with a neutral glutamine; (3) Tyr113Glu/Glu181Gln, to test if a glutamate at site 113 would be able to take on the role of a counterion if the counterion at site 181 is removed; and (4) Cys87Ser, a change from cysteine to serine in a nearby position (site

90) has been shown to longwave shift avian UVS pigment by 35 nm (Yokoyama et al., 2000), indicating that perhaps the same effect may occur for site 87.

It was hypothesised that if the UV sensitivity of tmt10 is due to an unprotonated Schiff base, then substituting the tyrosine at site 113 with a negatively charged glutamate may permit the binding of H^+ and shifts the λ_{max} to violet region, as observed in many non-SWS1 visual pigments (Sakmar et al., 1989, Terakita et al., 2004). However, spectral analysis showed that the Tyr113Glu mutants reconstituted with 11-*cis* and all-*trans* retinals remained UVS in their dark state, with λ_{max} at 368 nm (Fig. 5.3C) and 370 nm (Fig. 5.3D) respectively. This suggests that site 113 is not involved in spectral tuning of tmt10. The findings also indicated that tmt10 may already be protonated using other counterions, because the presence or absence of a negatively charged residue at site 113 does not have an effect on its spectral shift, despite an earlier (albeit unclear) suggestion that the pigments may be unprotonated.

Since the amino acid at site 113 does not seem to be involved in spectral tuning through its potential role as a counterion, it was deduced that perhaps Glu181 provides the counterion for protonation of the pigment. By substituting the glutamate residue at site 181 with a glutamine, this removed the negative charge present at this potential counterion site, which may result in a spectral shift towards shorter wavelengths. However, no pigments have ever been shown to have λ_{max} lower than 360 nm. The reconstituted Glu181Gln pigments with 11-*cis* as well as all-*trans* retinal chromophores failed to regenerate *in vitro*. This result suggests that Glu181 probably act as counterion to the tmt10 opsin, and its presence is vital for the stability of the structure as well as function of the pigment.

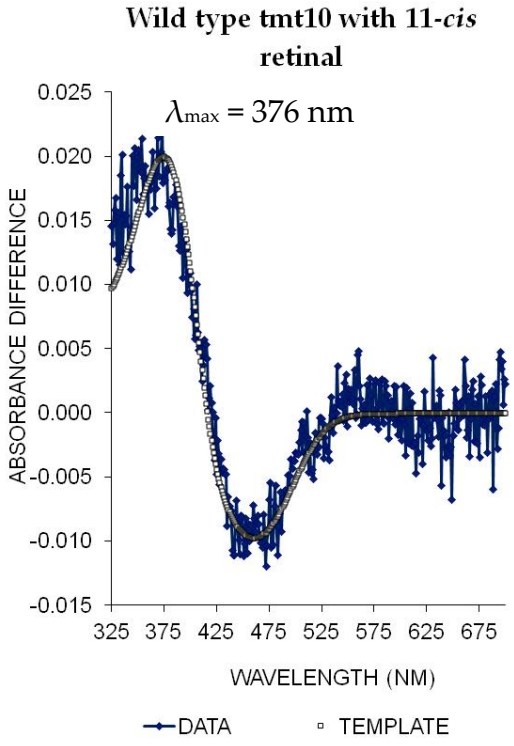
To confirm the importance of a counterion in stability of protonation of tmt10 opsin in the light, a double mutant carrying Tyr113Glu and Glu181Gln substitutions was generated. As the putative counterion at site 181 was removed by Glu181Gln substitution, replacement of the tyrosine with glutamate at site 113 might act as a counterion instead. This double mutant was successfully reconstituted *in vitro* with 11-*cis* and all-*trans* retinals, and the λ_{max} were measured to be 375 nm (Fig. 5.3E) and 367 nm (Fig. 5.3F) respectively. This finding suggests that site 113 is located in close proximity to the Schiff base, and adding in a counterion at site 113 when one is lacking at site 181 can help to regain protonation stability. Hence, the photopigment requires at least one negatively charged residue present in the vicinity of the Schiff base (e.g. at position 113 or 181) in order to maintain its tertiary structure. In addition, this result also supports the previous observation that wild-type tmt10 is already protonated in its dark state, because adding in a negative charge at position 113 in the Tyr113Glu/Glu181Gln mutant balances out the positive charge on the Schiff base and helps to retain UV-sensitivity of the pigment. If the tmt10 double mutant was unprotonated, addition of a counterion to the Schiff base at site 113 is predicted to give a spectral shift to the violet range. Since this did not occur, this again suggests that tmt10 double mutant was likely to be protonated.

Amino acids at position 90 and 93 have been implicated as one of the spectral tuning sites for avian VS and UVS pigments (Hunt et al., 2004, Odeen et al., 2011), whilst that at site 86 is important for non-avian SWS1 pigment (Cowing et al., 2002, Fasick et al., 2002, Deeb et al., 2003). It has been shown that cysteine is found in positions 90 and 93 of most of the avian UVS pigments studied, whereas serine is

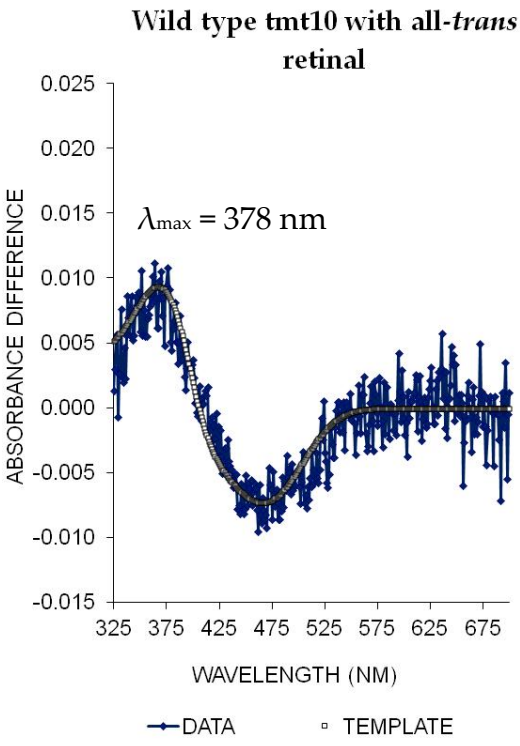
usually present at these sites in the VS pigments (Hunt et al., 2004, Odeen and Hastad, 2009). At site 87, it was noted that cysteine is present in all the zebrafish tmt opsins as well as SWS1 pigments. Since site 87 is located close to other sites (86, 90 and 93) that have been implicated for spectral tuning of UVS and VS pigments, the amino acid at this site was investigated. However, replacement of the cysteine with serine at a nearby site 87 in tmt10 did not cause a significant longwave shift from the wild-type tmt10. In the presence of 11-*cis* and all-*trans* retinals, the λ_{max} measured were 372 nm (Fig. 5.3G) and 378 nm (Fig. 5.3H) respectively. The result therefore suggests that site 87 does not influence the spectral tuning of the tmt10 pigment and other sites are likely to underpin its UVS.

Taken together, data from these mutagenesis experiments suggest that tmt10 is protonated in the wild-type pigment. Failure of the Glu181Gln mutant to regenerate demonstrated that the glutamate residue at site 181 is structurally vital, and may act as a counterion to counterbalance the protonated Schiff's base. However, in the absence of Glu181, replacement of the tyrosine at site 113 with a glutamate residue (i.e. Tyr113Glu/Glu181Gln mutant) can adopt the functional role as a counterion. Nonetheless, Glu181 seems to be the primary counterion used in this opsin, as the Tyr113Glu mutant is still shortwave shifted by 8 nm with 11-*cis* retinal and 15 nm with all-*trans* retinal compared to the wild-type tmt10. These results also support the previous spectrophotometric data in that tmt10 functions as a bistable pigment, because most of the mutants generated can form stable interactions with 11-*cis* as well as all-*trans* retinals.

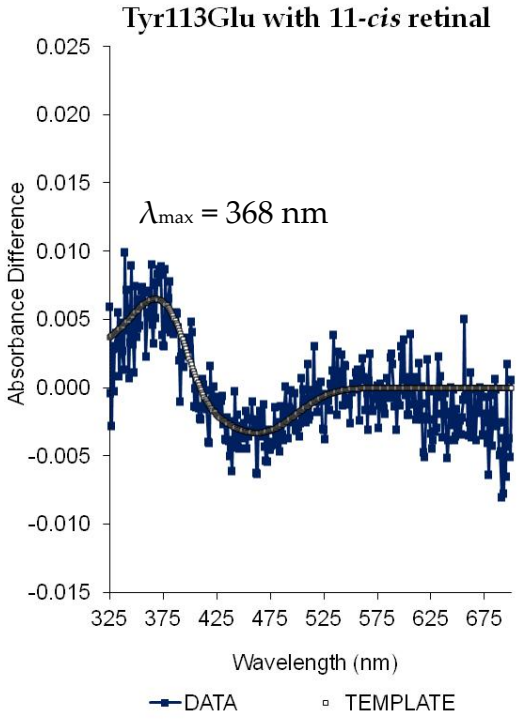
(A)


Wild type tmt10 with 11-cis retinal
λmax = 376 nm
ABSORBANCE DIFFERENCE
0.025
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
325 375 425 475 525 575 625 675
WAVELENGTH (NM)
DATA
TEMPLATE

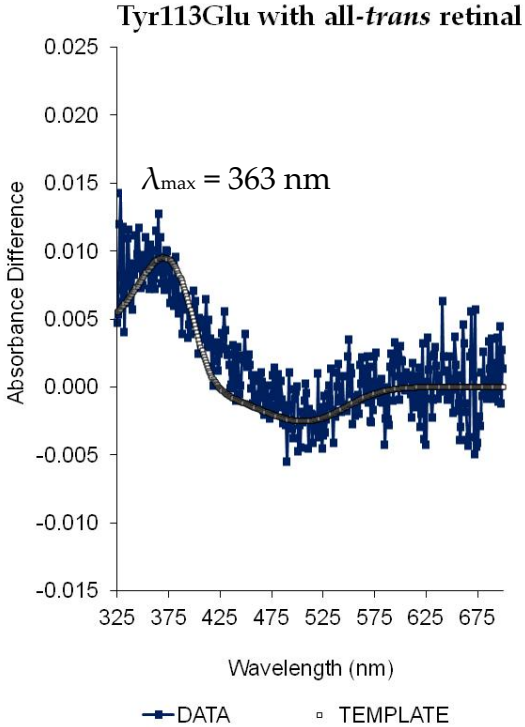
(B)


Wild type tmt10 with all-trans retinal
λmax = 378 nm
ABSORBANCE DIFFERENCE
0.025
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
325 375 425 475 525 575 625 675
WAVELENGTH (NM)
DATA
TEMPLATE

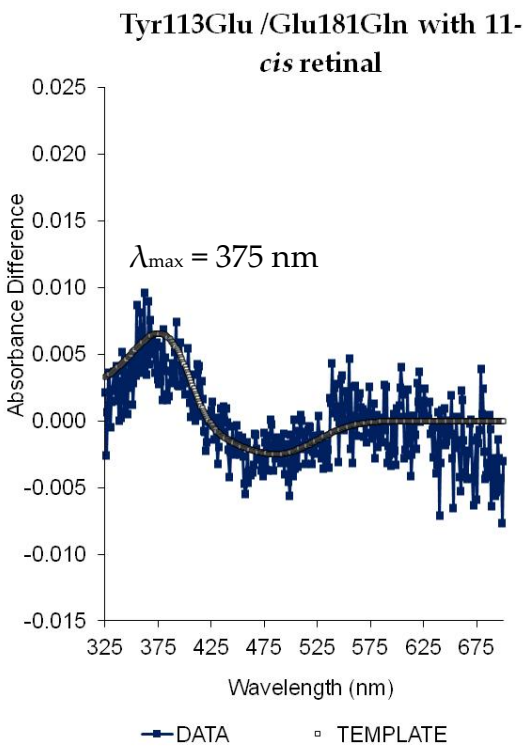
(C)


Tyr113Glu with 11-cis retinal
λmax = 368 nm
Absorbance Difference
0.025
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
325 375 425 475 525 575 625 675
Wavelength (nm)
DATA
TEMPLATE

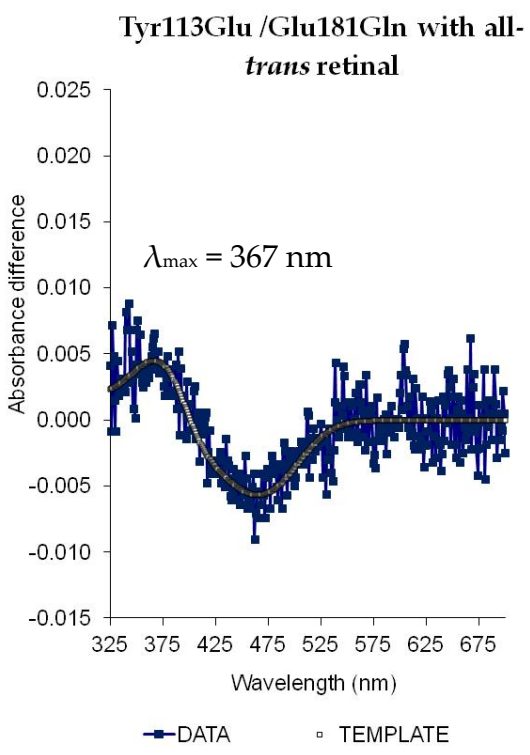
(D)


Tyr113Glu with all-trans retinal
λmax = 363 nm
Absorbance Difference
0.025
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
325 375 425 475 525 575 625 675
Wavelength (nm)
DATA
TEMPLATE

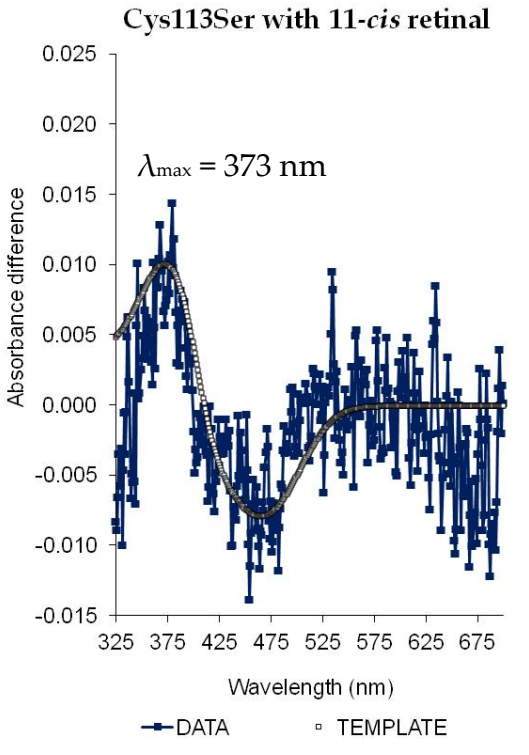
(E)


Tyr113Glu /Glu181Gln with 11-cis retinal
λmax = 375 nm
Absorbance Difference
Wavelength (nm)
DATA
TEMPLATE

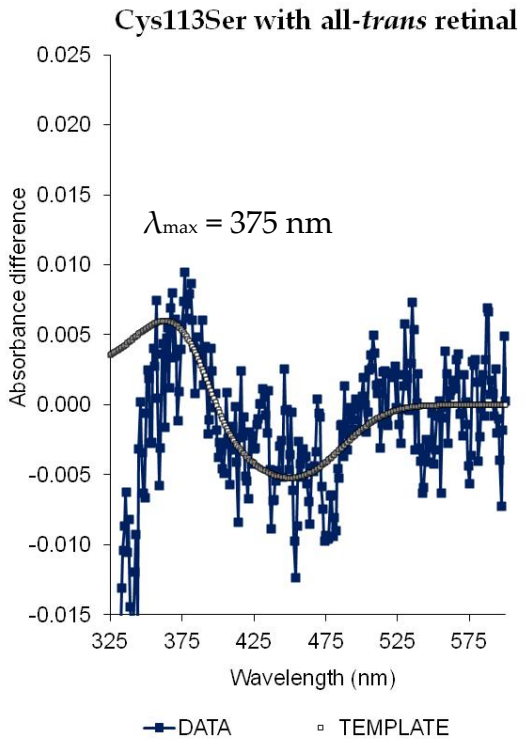
(F)


Tyr113Glu /Glu181Gln with all-trans retinal
λmax = 367 nm
Absorbance difference
Wavelength (nm)
DATA
TEMPLATE

(G)


Cys113Ser with 11-cis retinal
λmax = 373 nm
Absorbance difference
Wavelength (nm)
DATA
TEMPLATE

(H)


Cys113Ser with all-trans retinal
λmax = 375 nm
Absorbance difference
Wavelength (nm)
DATA
TEMPLATE

(I)

tmt10 variants	Predicted λ_{max} region		Measured λ_{max} with 11-*cis* retinal	Measured λ_{max} with all-*trans* retinal
	If protonated	If unprotonated		
Wild-type	UV	UV	376 nm	378 nm
Tyr113Glu	UV	Violet	368 nm	363 nm
Glu181Gln	UV?	UV?	Failed	Failed
Tyr113Glu/Glu181Gln	Violet	Violet	375 nm	367 nm
Cys87Ser	Violet	Violet	373 nm	375 nm

Figure 5.3 Difference spectra for mutant tmt10 opsins regenerated *in vitro*. Tyr113Glu, Glu181Gln, Tyr113Glu/Glu181Gln, and Cys87Ser were generated by site-directed mutagenesis of the wild type tmt10 cDNA sequence. The wild type and mutant of tmt10 were expressed in HEK293T cells, harvested and regenerated as described in Section 2.5.3.2 and 2.5.3.4. The graphs show the difference spectra of **(A, B)** wild-type tmt10, **(C, D)** Tyr113Glu, **(E, F)** Tyr113Glu/Glu181Gln, and **(G, H)** Cys87Ser reconstituted with 11-*cis* and all-*trans* retinals. For each pigment, difference spectra were obtained by subtracting the acid-denatured spectra from the dark spectra. The difference spectra were then fitted with a modified Govardovskii template (Govardovskii et al., 2000, Parry et al., 2004) for determination of λ_{max} values. **(I)** A table summarising the predictions of λ_{max} region for protonated or unprotonated pigments, as well as the measured values of λ_{max} with 11-*cis* and all-*trans* retinals from UV-Vis spectrophotometric analysis.

5.3 Discussion

In this study, zebrafish tmt opsins were expressed in HEK293T cells and were successfully reconstituted with 11-*cis* and all-*trans* retinoids. Characterisation of the regenerated tmt pigments by UV-Vis spectrophotometric analysis demonstrated that they are a group of UVS bistable photopigments. That is, these opsins can form stable interactions with both 11-*cis* and all-*trans* retinal chromophores. Interestingly, these two forms of photoisomers are both UV-light absorbing. The ability to bind two (or more) forms of retinal chromophore is a defining characteristic of bistable pigments, which has also been observed in other non-visual opsins such as opn4, opn5, and parapinopsin (Mure et al., 2009, Yamashita et al., 2010, Koyanagi et al., 2004). Photopigments that are monostable, e.g. rod and cone opsins and some non-visual opsin classes (e.g. VA, pinopsin, parietopsin, some forms of opn4) (Okano et al., 1994, Davies et al., 2011, Davies et al., 2012b, Sakai et al., 2012), usually only bind 11-*cis* retinal (Kefalov et al., 2001). A recent paper has shown that a pufferfish tmt pigment exhibited λ_{max} at ~460 nm in the dark (Koyanagi et al., 2013), which is significantly longwave shifted from the λ_{max} of zebrafish tmt pigments. Comparison of known spectral tuning sites in the pufferfish tmt pigment with those of zebrafish tmt pigments and other vertebrate UVS pigments indicates that Leu52 might be involved in driving violet sensitivity (see Table 5.2). It is also likely that there are other unidentified tuning sites that may contribute to the spectral differences observed. In addition, Koyanagi et al. (2013) demonstrated that mouse Opn3 absorbs light in the visible region with λ_{max} at ~500 when reconstituted with 9-*cis* and 11-*cis* retinals. This

Table 5.2 *Amino acid residues at sites implicated for spectral tuning in UVS pigments of teleost fish, amphibia, reptiles, birds, and mammals.*

			Tuning sites							
Class	**Species**	**Pigments**	**46**	**49**	**52**	**86**	**90**	**93**	**114**	**118**
Teleost fish	Zebrafish (*Danio rerio*)	tmt2l	Phe	Thr	Phe	Val	Gly	Phe	Gly	Ser
		tmt2s	Phe	Thr	Phe	Val	Gly	Phe	Gly	Ser
		tmt6	Ser	Ile	Thr	Val	Gly	Leu	Gly	Ser
		tmt9	Phe	Thr	Phe	Val	Gly	Leu	Gly	Ser
		tmt10	Ile	Val	Phe	Val	Gly	Leu	Gly	Ala
		tmt14	Ala	Leu	Cys	Val	Gly	Phe	Gly	Ser
		tmt24	Phe	Thr	Phe	Val	Gly	Phe	Gly	Ser
	Pufferfish (*Takifugu rubripes*)	tmt	Phe	Thr	Leu	Val	Gly	Phe	Gly	Ser
	Tramitichromis intermedius	VS	Phe	Trp	Thr	Val	Gly	Thr	Gly	Thr
	New Yellow Regal Peacock (*Aulonocara baenschi)*	VS	Phe	Trp	Thr	Val	Gly	Thr	Gly	Thr
Amphibia	Salamander *(Ambystoma tigrinum)*	UVS	Phe	Phe	Thr	Phe	Ser	Thr	Ala	Ser
Reptiles	Chameleon (*Anolis carolinensis)*	UVS	Phe	Phe	Thr	Phe	Ser	Thr	Ala	Ser
	Clawed frog (*Xenopus laevis)*	VS	Met	Phe	Thr	Met	Ser	Pro	Ala	Thr
Birds	Budgerigar (*Melopsittacus undulatus*)	UVS	Phe	Met	Thr	Ala	Cys	Thr	Gly	Ala
	Canary (*Serinus canaria*)	UVS	Leu	Val	Thr	Cys	Cys	Thr	Gly	Ala
	Zebra finch (*Taeniopygia guttata*)	UVS	Leu	Val	Thr	Cys	Cys	Thr	Gly	Ala
	Pigeon *(Columba livia)*	VS	Phe	Phe	Thr	Ser	Ser	Thr	Ala	Ala
	Chicken (*Gallus gallus*)	VS	Ile	Phe	Thr	Ser	Ser	Val	Ala	Thr
Methatherian mammals	Dunnart *(Sminthopsis crassicaudata)*	UVS	Phe	Phe	Thr	Phe	Ser	Thr	Gly	Ser
Eutherian mammals	Mouse *(Mus musculus)*	UVS	Phe	Phe	Thr	Phe	Ser	Thr	Ala	Ser
	Cow *(Bos taurus)*	VS	Phe	Phe	Thr	Tyr	Ser	Ile	Ala	Cys

Sequence accession numbers: pufferfish (AF402774), *Tramitichromis intermedius* (GQ422529), New Yellow Regal Peacock (GQ422528), salamander (AF038948), chameleon (AF134194), budgerigar (Y11787), canary (AJ277922), zebra finch (AF222331), pigeon (AJ277991), chicken (M92039), Dunnart (AY442173), mouse (AF190671), cow (U92557).

NB: Amino acid substitutions at site 46, 49, 52, 86, 93, 114 and 118 are correlated with UV/violet shifts in non-avian SWS1 pigments (Shi et al., 2001), whilst site 90 and 93 were found to be important for avian UVS pigments (Wilkie et al., 2000).

was not surprising as Opn3 possesses a negatively charged Asp113, which can act as a counterion to the Schiff base linkage. A negatively charged residue of Glu113 is observed in many non-SWS1 visual pigments that are violet sensitive (Sakmar et al., 1989, Terakita et al., 2004), and is absent in the UVS zebrafish tmt opsins.

To date, all UVS pigments discovered can be divided into four groups: (1) invertebrate visual opsins, e.g. honey bee UVS pigments (Townson et al., 1998); (2) a subset of *SWS1*-encoded vertebrate visual opsins, e.g. mouse S-cones (Lyubarsky et al., 1999), marsupial SWS1 pigments (Palacios et al., 2010); (3) vertebrate parapinopsin (Wada et al., 2012); and (4) vertebrate opn5 (Yamashita et al., 2010). From the phylogenetic analysis in Chapter 3, it seems that these four groups of opsins and tmt opsins have evolved independently to absorb UV light. Examination of the potential counterion sites of a representative tmt opsin (tmt10) by site-directed mutagenesis demonstrated that like many other non-visual opsins, tmt10 employs Glu181 as its counterion for maintaining the stability of photopigment at dark state. The failure of Glu181Gln tmt10 mutant to regenerate *in vitro* indicated that the presence of a glutamate at this site is critical for the structure as well as the subsequent function of the photopigment. However, in the absence of Glu181, replacement of the tyrosine at site 113 with a glutamate can adopt this role. These findings, together with the phylogenetic positioning of tmt opsins as ancestral opsins predating VA and visual opsins, strongly support the hypothesis that counterion displacement from Glu181 to Glu113 might have occurred during the course of evolution in other contemporary opsins (Terakita et al., 2004). Terakita et al. (2004)

suggests that parietopsin is the location of this displacement, but tmt opsins indicate that was occurred earlier.

Results from the mutagenesis study and the pH-dependency experiments with tmt10 suggest that tmt pigments may employ a novel spectral tuning mechanism that has not yet been reported for UV-light sensitivity. Site-directed mutagenesis testing site 87, which is in close proximity to a spectral tuning site 86 in non-avian SWS1 pigments, demonstrated that this site is not involved in mediating UV-sensitivity of tmt10. Other spectral tuning sites such as site 90, 93, 114, 116, and 118 will need to be tested for a full comparison with known UVS pigments (Shi and Yokoyama, 2003). When regenerated under different pH conditions, tmt10 did not exhibit a spectral shift and maintained UVS. Collective data suggest that the Schiff base is either already protonated at the dark state, or that protonation is not important and thus the Schiff base is constitutively unprotonated. All of the known UVS pigments reported possess properties indicative of an unprotonated retinyl Schiff base in the dark state (Kono, 2006, Babu et al., 2001). It is thought that upon photoactivation, protonation and subsequent de-protonation of this Schiff base are critical for the generation and degradation of intermediates during the phototransduction process (Dukkipati et al., 2001, Dukkipati et al., 2002). However, several lines of evidences from this study have shown that tmt10 pigment is already protonated in its dark state, implicating a different mechanism for UV absorbance by UV-light.

Since both 11-*cis* and all-*trans* retinal-binding forms of tmt pigments exhibit a λ_{max} in the UV region, this bistable property suggests that steady UV-light irradiation

may cause continuous activation of the photopigments as the isomers interconvert between the two forms. Other bistable UVS opsins (e.g. opn5 and parapinopsin) are usually changed into visible light-absorbing photoproducts (Yamashita et al., 2010, Wada et al., 2012). Steady UV-irradiation would therefore be predicted to result in a certain degree of suppression in the phototransduction cascade of these non-tmt opsins, as the resulting photoproducts can only be reverted back to the dark state by visible light. It has been postulated that this shift in λ_{max} of the photoproduct is due to conformational changes within the pigment structure, causing the release of all-*trans* retinal before 11-*cis* retinal binds to the opsin again (Imai et al., 1997). Unlike these reported UVS pigments, tmt pigments regenerated with 11-*cis* and all-*trans* retinal chromophores are both UVS, indicating that the photoreaction between tmt pigments and their photoproducts is likely to be reversible under UV irradiation. It is possible this unique ability to maintain a constant level of active tmt photopigment in response to the damaging UV-light serves as a protective mechanism for the organisms. UV irradiation can lead to genetic changes in the DNA, and it may be that tmt opsins are coupled to DNA repair enzymes such as photolyase to trigger their repair mechanism upon light activation.

CHAPTER 6

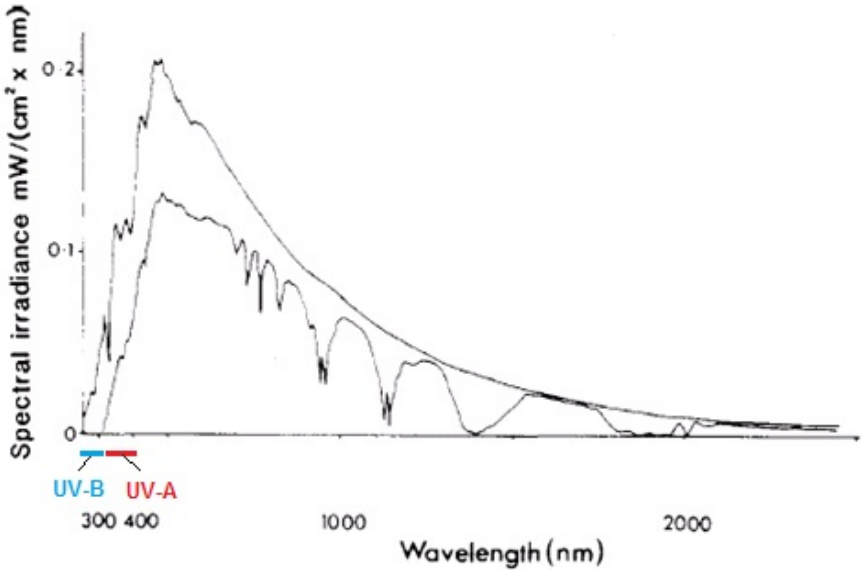
Potential Signalling Cascades Induced by Activation of tmt Opsins

CHAPTER 6: Potential Signaling Cascades Induced by Activation of tmt Opsins

6.1 Introduction

Sunlight contains wavelengths in the ultraviolet (UV) range of the electromagnetic spectrum, consisting of about 8.5% UV-A radiation (320-400 nm) and 0.3% UV-B radiation (280-320 nm) (Moseley, 1988, 2002) (see Fig. 6.1A). Whilst the majority of UV-B is absorbed by ozone (particularly in the stratosphere layer), UV-A is not absorbed by the atmosphere and penetrates through to terrestrial and aquatic environments (Moseley, 1988). Although UV radiation comprises only about 8% of the total solar radiation reaching the inner atmosphere and about 50% of this is filtered out at sea level, it still plays a crucial role in a variety of biological processes such as vision, species recognition, orientation and navigation in fish (Cronin et al., 2003, Britt et al., 2001). It has been demonstrated that in shallow Indian rivers where zebrafish inhabits, there is a transmittance level of ~60% for short-wavelength light at ~380 nm (Fig. 6.1B). Therefore, organisms also have to be protected from the harmful effects of UV radiation as overexposure can cause strand breakage and mutations in DNA, leading to changes in cell cycle and the potential development of cancer (Griffiths et al., 1998). For example, humans have UV filtering compounds in the lens that absorb in the UV range (<400 nm) and protect the retina from UV-induced damages (Wood and Truscott, 1993, Bova et al., 1999). A variety of species have also evolved DNA repair mechanisms, such as nucleotide excision repair for single-strand damage (Sancar and Reardon, 2004) and recombinational repair for double-strand breakage (Cox, 1999, Smith, 2004) to overcome the mutations induced by UV.

(A)



(B)

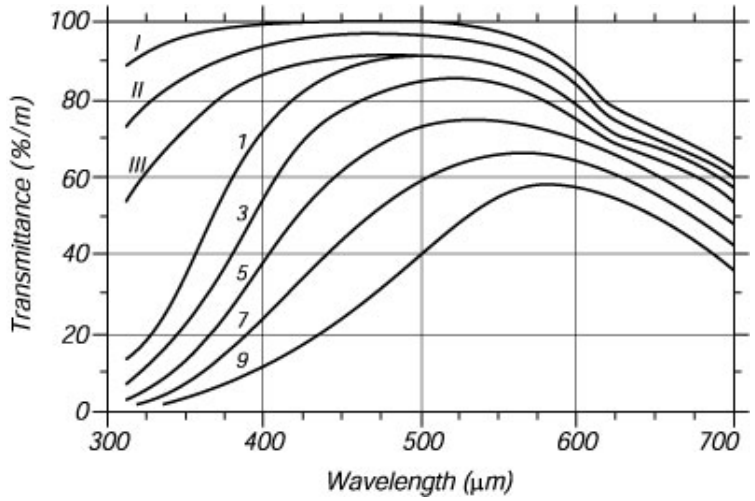


Figure 6.1 (A) Irradiance spectrum of the sun measured from outside the atmosphere of the Earth (upper curve) and from sea level (bottom curve) (Moseley, 1988). **(B)** The effects of coastal water and ocean water on the filtering properties of light at wavelengths of 300 to 700 nm (Jerlov, 1976). The optical properties of the water are classified as follow: type 1-9 for coastal water, corresponds to increasing turbidity of the water; and type I-III for ocean water, with type I being the clearest water, type II as turbid tropical-subtropical water, and type III as the most turbid oceanic water (Jerlov, 1976). The habitat of zebrafish may be assigned to type 1, which is the least turbid, shallow coastal water.

This study has identified a group of UV-sensitive (UVS) tmt opsins in zebrafish, which have a widespread pattern of expression in many of their tissues and organs. UV-vis spectrophotometric analysis demonstrated that tmt photopigments can form stable interactions with both 11-*cis* and all-*trans* retinal *in vitro* (see Chapter 5) and absorbed with spectral peaks in the UV region at around 370 nm (i.e. UV-A). In order to decipher the functional roles that UVS tmt pigments may have, it is important to determine the signalling pathway(s) utilised by each of the tmt opsins. Although UVS pigments (e.g. SWS1 opsins) have already been identified in zebrafish and other teleosts, they appear to be limited to a subset of cone photoreceptors (Allison et al., 2010, Raymond and Barthel, 2004) and the pineal gland (Forsell et al., 2001). The newly identified UVS tmt pigments are more ubiquitously expressed, as they were detected in nearly all peripheral and CNS tissues of zebrafish (Chapter 4). Their discovery may therefore represent the molecular basis for broad UV reception in teleosts.

Within the opsin superfamily, the majority of pigments generally couple to one of five different G-proteins, e.g. G_t, $G_{i/o}$, G_s, G_q, or G_v-type G proteins, leading to the activation of an intracellular transduction cascades (see Fig. 1.5) (Terakita, 2005, Oka et al., 2009). A small number of opsins, including retinal G-protein-coupled receptor (RGR) opsin, squid retinochrome, and peropsin (RRH), do not bind any G-proteins. Instead, they appear to function as photoisomerases that convert all-*trans* retinal into its 11-*cis* form upon irradiation with orange light (>520 nm) (Hara and Hara, 1976, Shen et al., 1994, Terakita et al., 2000, Koyanagi et al., 2002). The photoisomerase characteristic of these opsins is thought to be important for the

regeneration and trafficking of 11-*cis* retinal chromophore to be used by other opsins that do couple to G proteins (Maeda et al., 2003, Chen et al., 2001). Phylogenetic analysis and the genomic organization of zebrafish *tmt* opsins suggest that they are more closely related to opsins which activate G-proteins (e.g. visual opsins) than to the photoisomerases (e.g. RRH and RGR opsin) (Chapter 3). However, it remains to be ascertained whether zebrafish tmt opsins act as light-sensitive GPCR photopigments or photoisomerase molecules.

In this chapter, the aim was to determine the G-protein signalling cascades activated by zebrafish tmt opsins using various functional assays following transfection into mammalian cell lines, including: (1) whole-cell patch clamp recording (Melyan et al., 2005), as used previously to study functional properties of numerous opsin photopigments (Pires et al., 2009, Davies et al., 2011); (2) *in vitro* Ca^{2+} imaging (Panda et al., 2005, Kumbalasiri et al., 2007, Hughes, unpublished), for testing of the G_q-signalling pathway; and (3) a bioluminescence reporter assay (Bailes et al., 2012), for detection of G_q, G_s and G_i-activity. In an effort to study the function of tmt opsin in an intact cellular environment, this project employed mammalian heterologous expression systems to examine the photoresponses induced. The well characterized murine neuroblastoma (Neuro-2A) and Human Embryonic Kidney 293T (HEK-293T) cell lines were selected for these studies, based on their proven ability to support opsin receptor G-protein signalling (Kumbalasiri et al., 2007, Spencer et al., 1997, Melyan et al., 2005, Pires et al., 2009). However, it is important to note that there are potential problems using these non-native heterologous cell lines, as opsins can promiscuously bind different G proteins. Identifying a particular

signalling pathway in the model system therefore may not be a true representation of the signalling cascades *in vivo*. A closer approach would be to use a zebrafish cell line which do not possess inherent photoentrainable clocks (e.g. ZF-13; Whitmore, unpublished data), though, none have been characterised for opsin expression studies. In addition, all current zebrafish cell lines are known to express several of the zebrafish opsins, and tmt opsin itself has already been identified in the PAC-2 cells (Moutsaki et al., 2003). Hence, these cell lines are already light responsive and are not suitable for studying light sensing function of tmt opsins here. For this reason, Neuro-2A and HEK-293T were chosen as the most appropriate model systems for this research. From the results, potential signalling cascades of zebrafish tmt opsins will be discussed.

6.2 Results

6.2.1 Immunocytochemistry (ICC) confirms tmt expression in Neuro-2A cells

Immunocytochemistry (ICC) analysis was performed to verify the heterologous expression of zebrafish tmt opsins in a mouse (Neuro-2A) cell line, and the trafficking of the opsin proteins to plasma membrane. Plasmids of 1D4-tagged tmt2s, tmt2l, tmt6, tmt9, tmt10, tmt14, and tmt24 were transfected into Neuro-2A cells as described in Section 2.5.2. Expression of each tmt opsin protein was tracked by immunolabeling using a primary mouse anti-1D4 monoclonal antibody (Molday and MacKenzie, 1983) that was diluted 1:2000 and a secondary Alexa 488 donkey anti-mouse IgG antibody (Invitrogen, UK) that was diluted 1:500 (see Section 2.5.4.1 for details). Immunocytochemical results demonstrated that labelling of all seven

tmt opsins was detectable at the plasma membrane of both the cell body and the cellular processes of transfected Neuro-2A cells, as expected for GPCRs (Fig. 6.2). They were also observed in the cytoplasm and possibly the endoplasmic reticulum (ER), which is common with transient over-expression of plasmids in cells by liposome-based transfection. No labelling was observed in (untransfected) cells (Fig. 6.2H). Morphologies of the transfected cells generally varied significantly, with different cell size and length of processes. However, repeated experiments (n=4) consistently showed that cells expressing tmt2s developed a distinctive morphology, as they lacked processes and exhibited rounder cell bodies (Fig. 6.2G). Based on observation, it is postulated that the morphology of *tmt2*s-transfected cells might be correlated with higher cell density. This effect of *tmt2* expression in darkness and absence of retinal suggests a constitutively active component to the tmt2 proteins following their expression in cells. Further experiments will need to be performed to evaluate if expression of tmt2 increases the proliferation rate of Neuro-2A cells. Nonetheless, ICC results showed that all opsins seem to be correctly folded for insertion in the ER membrane for trafficking to the membrane of Neuro-2A cells.

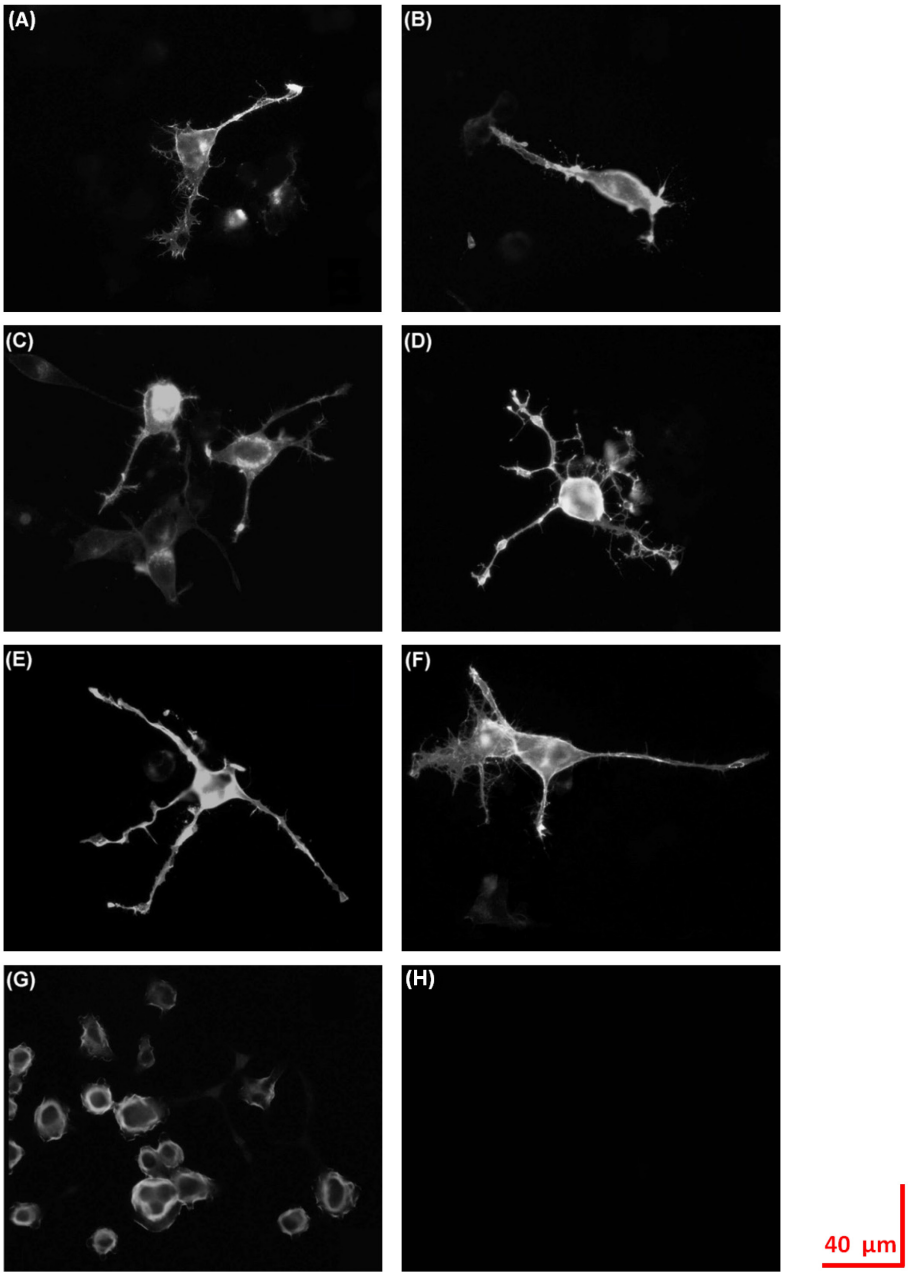


Figure 6.2 Immunocytochemical analysis of Neuro-2A cells transfected with **(A)** *tmt9*, **(B)** *tmt6*, **(C)** *tmt10*, **(D)** *tmt14*, **(E)** *tmt24*, **(F)** *tmt2l* and **(G)** *tmt2s*, all of which are 1D4 tagged constructs. The cells were stained with a primary mouse anti-1D4 antibody then with a secondary Alexa 488 donkey anti-mouse IgG antibody. **(H)** For untransfected Neuro-2A cells that were used as negative controls, no staining was observed above background. Images were taken using a fluorescence Olympus IX71 inverted microscope in black and white for enhanced contrast. Scale bars in red denoting 40 µm are shown in right bottom corner, and are representative for images A-H.

6.2.2 Calcium (Ca2+) imaging of Neuro-2A cells expressing tmt opsins

Monitoring of intracellular calcium concentration $[Ca^{2+}]_i$ is commonly used to investigate GPCR function, most notably $G_{q/11}$-type G-protein. The $G_{q/11}$ G-protein activates phospholipase C-beta (PLC-β), which is coupled to production of inositol trisphosphate (IP_3) and mediates the release of Ca^{2+} from the ER (Billups et al., 2006) (see Fig. 6.3). A Ca^{2+} imaging system has been set up in this laboratory to screen for pigment photosensitivity when expressed in Neuro-2A cells with retinal chromophore, in particularly those which are coupled to G_q (e.g. melanopsin) (Hughes et al., unpublished). Other $G\alpha$ subunits are also known to be expressed in the Neuro-2A cells but as the output measured is Ca^{2+} flux, this assay will predominantly report $G_{q/11}$-based responses. Human OPN4 (hOPN4) was included as a positive control, as previous studies have shown that it leads to photoresponses in a similar system (Kumbalasiri et al., 2007).

The responses triggered from tmt opsins were compared to hOPN4 by measuring Ca^{2+} transient traces in Neuro-2A cells with rhodamine-2 acetoxymethyl (Rhod-2 AM). Rhod-2AM is a cationic indicator dye used for detecting changes in cytosolic and mitochondrial Ca^{2+} concentration. In a Ca^{2+}-bound form, excitation by light at 552 nm drives the emission of a fluorescent signal at 581 nm (Invitrogen, UK). The main reason for using Rhod-2AM dye in this experiment is that excitation of the dye with 545 nm light would permit the usage of lower wavelengths (e.g. 420 and 480 nm) as second light pulse stimulations in the existing scope setup. Although the spectrophotometric data in Chapter 5 demonstrated that the λ_{max} of tmt opsins are in the UV-range between 370-380 nm, 420 and 480 nm light were selected for

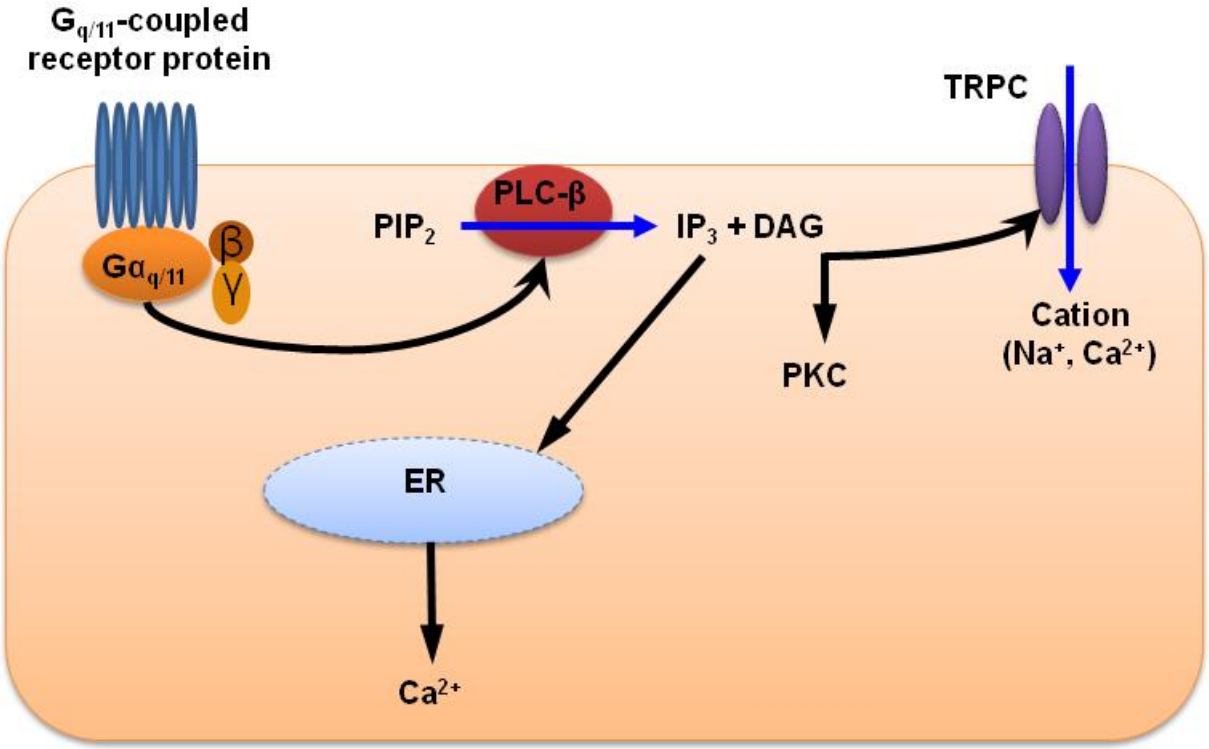


Figure 6.3 The signalling pathway of a $G_{q/11}$-type G protein. Activation of a $G_{q/11}$-coupled receptor protein stimulates the activity of phospholipase C (PLC-β). PLC-β hydrolyses membrane-bound phospholipid (PIP_2) to produce two secondary messengers: (1) inositol-1,4,5-phosphate (IP_3), and (2) diacylglycerol (DAG). IP_3 binds to IP_3 receptor in the membrane of endoplasmic reticulum (ER) and triggers the release of intracellular Ca^{2+}, whilst DAG activates protein kinase C (PKC) and possibly canonical transient receptor potential (TRPCs) channels which are permeable to cations (e.g. Na^+ and Ca^{2+}). (Diagram was modified from Hughes et al., 2012 and Lehninger et al., 2013)

stimulation instead. This is because UVA-irradiation has been shown to have phototoxic effects on cells, which include generation of free radicals that leads to cytoplasmic damage and causes influx of Ca^{2+} into the cells (Mendez and Penner, 1998, Kimura et al., 1992). The 420 and 480 nm light chosen were predicted to be within the absorption spectral range of the tmt opsins, which should be sufficient to activate the pigments for functional experiments. Hughes et al. (unpublished) have shown that melanopsin can be activated by 540 nm light, which is 60 nm past the reported 480 nm peak (Melyan et al., 2005). The experiments were performed with or without chromophore (9-*cis* or all-*trans* retinal) at room temperature (see Section 6.5.2 for detail of methodology). These two retinals were selected to test for bistablity of the pigment molecules, as most opsins are known to be coupled with 11-*cis* and/or all-*trans* retinal (Section 1.3.4 and 1.3.5). However, instead of 11-*cis* retinal, its analogues 9-*cis* retinal was used for probing the functions of tmt opsins since 9-*cis* retinal can also bind apo-opsins to form active photopigments (Pepperberg et al., 1978, Sekharan et al., 2007). Images were initially collected every 2 s using a 100 ms pulse of 545 nm for 1 min (image collection pulses), then they were interspaced with fifteen 1 s light pulse stimulation at either 420 or 480 nm over 30 s (as illustrated by the bars under each graph in Fig. 6.5A and 6.6A).

The positive (hOPN4-transfected) and negative (untransfected) controls, as well as the tmt-transfected cells with no chromophore in each set of experiments confirmed that the methods of transfection and dye loading were successful. Dye signals were reported as a relative change in fluorescence intensity ($dF/F = (F_t - F_{t0})/F_{t0}$), where F_t is the fluorescence signal measured at a given time and F_{t0} is the

baseline fluorescence at dt = 0. In the presence of 9-*cis* retinal, positive control cells exhibited a rapid rise in intracellular Ca^{2+} in response to a set of imaging protocol and exposure to 545 nm image collection pulses, the mean amplitudes of dF/F values were 0.4-0.6 (Fig. 6.5A, 6.6A). The initial elevation in Ca^{2+} concentration $[Ca^{2+}]_i$ decreased to near baseline within 1 min. Stimulation with light of 420 or 480 nm that followed also triggered transient Ca^{2+} signals in the hOPN4-expressing cells, but the characteristics of these responses differed significantly between the two wavelengths: 480 nm light triggered rapid Ca^{2+} transient responses similar to that of the 545 nm, although their mean amplitudes were about 50% less than the initial response, likely due to desensitisation (Fig. 6.5, 6.7). In contrast, 420 nm light evoked a higher rise in $[Ca^{2+}]_i$, with mean amplitudes about 40% greater than the initial response to 545 nm and this increase in $[Ca^{2+}]_i$ was maintained for a prolonged period of time (Fig. 6.6 & 6.7). It is possible that this 420 nm light-evoked response resulted from the phototoxic effects of excess 9-*cis* retinal in the cells (Hughes et al., unpublished). Studies have shown that unbound all-*trans* retinal, which is converted from the 9-*cis* retinal upon irradiation, can generate retinal proteins (e.g. N-retinylidene-N-retinyl ethanolamine, A2E) that perturb cellular membrane integrity (Sparrow et al., 1999, Schutt et al., 2000). Such damage in the cell membrane can permit entry of Ca^{2+} ions, which triggers massive release of Ca^{2+} from the intracellular store, e.g. ER. This process, known as Ca^{2+}-induced Ca^{2+} response, might be the 420 nm responses detected here. Although the absorption maximum of all-*trans* retinal has been shown to be 380 nm (Sun and Nathans, 2001), the long-wave limb of the spectra displays a significant absorption level at 420 nm, much higher than at 480 nm (see Fig. 6.4). This

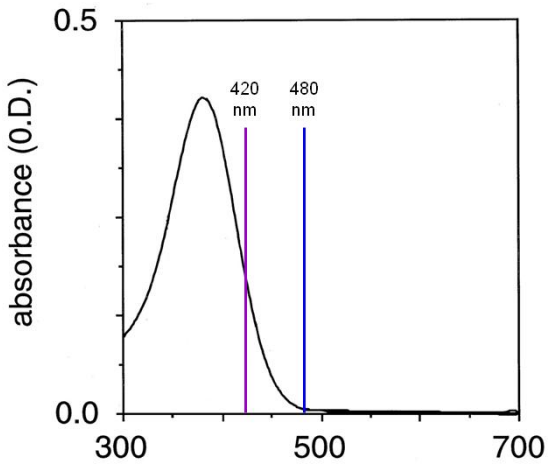


Fig. 6.4 Absorption spectra of all-*trans* retinal. The peak absorption value was measured at 380 nm. Diagram modified from Sun and Nathans, 2001.

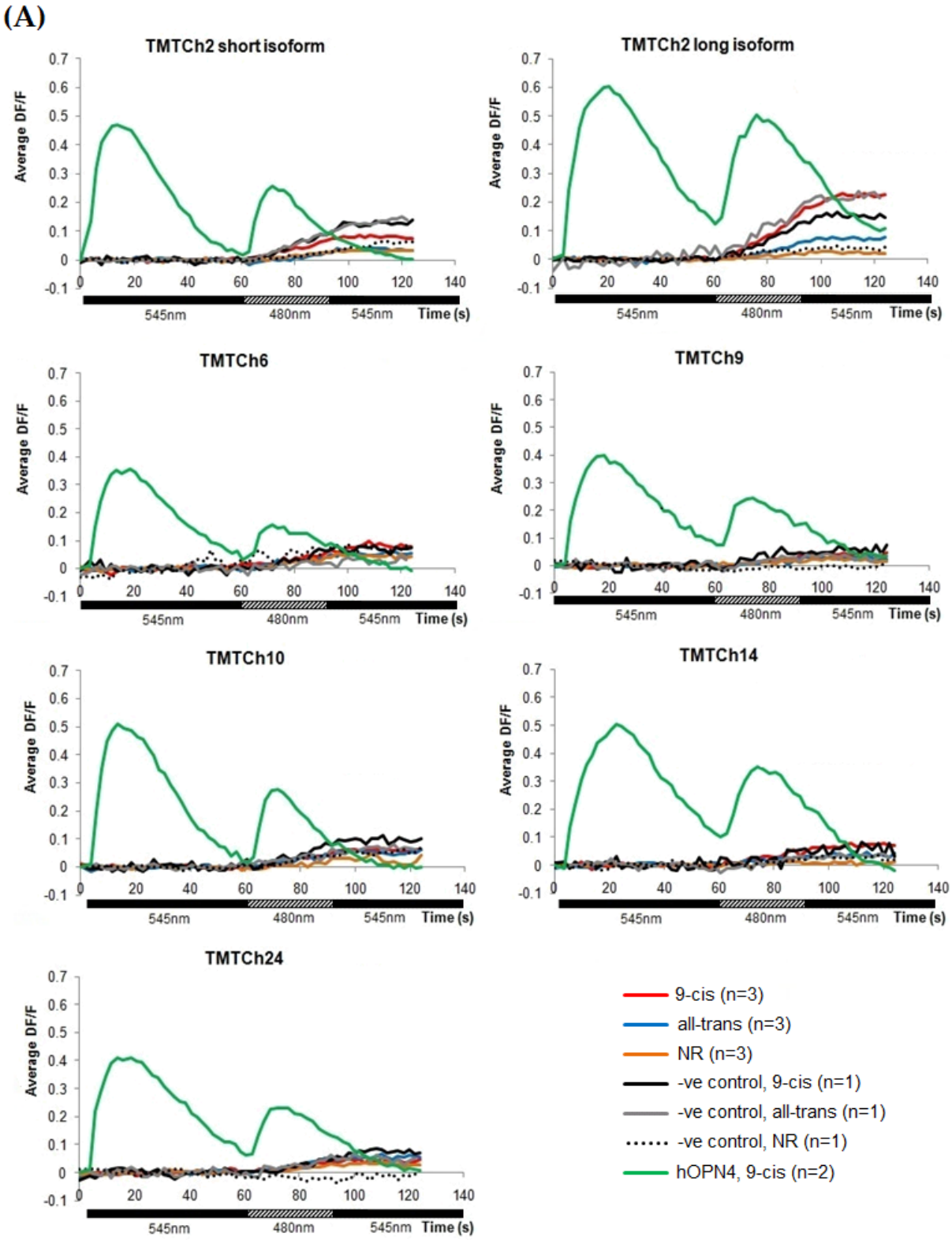
suggests that irradiation at 420 nm is likely to produce a significant level of activated all-*trans* retinal in the cells, which may contribute to light-induced damages in the cell membrane, driving the rapid Ca^{2+} response observed.

As for the negative control cells, the cytosolic Ca^{2+} levels were not affected by 545 nm light at all and the dF/F values remained at baseline level. After light stimulation at 480 nm, these cells exhibited a small increase in $[Ca^{2+}]_i$ with slow kinetics, and the effect seemed to be slightly more prominent in the presence of retinal with dF/F values reaching ~0.1 (Fig. 6.5). This could be caused by the low level of photo-excitation of unbound all-*trans* retinal that was already present or converted from 9-*cis* retinal. When subjected to 420 nm light exposure, cells exhibited a pronounced rise in $[Ca^{2+}]_i$ similar to the response observed in hOPN4-expressing cells but in the presence of retinal only (Fig. 6.6). The heightened level of $[Ca^{2+}]_i$ was also sustained for a long time in these cells. This 420 nm response observed can be explained by the Ca^{2+}-induced Ca^{2+} response described before, possibly driven by

photoactivation of free all-*trans* retinal. These data indicate that the 545 and 480 nm responses were specific to the hOPN4 transiently expressed in Neuro-2A cells. However, the robust 420 nm responses observed were unlikely to be opsin-specific and appeared to be more retinal-dependent.

Ca^{2+} imaging data recorded from cells transfected with each tmt opsin showed no significant difference to negative controls when stimulated by light at 545, 480 and 420 nm (Fig. 6.5B, 6.6B, and 6.7B). The slow rise of $[Ca^{2+}]_i$ in untransfected and *tmt*-transfected cells triggered by 480 nm light could be due to endogenous proteins (including other opsins) binding to retinal in the membrane and therefore photosensitising the Neuro-2A cells. However, these small changes in the Ca^{2+} signals observed might also be photopigment-independent. When compared with positive controls, differences in response kinetics to both 480 and 420 nm light might at least suggest that, unlike hOPN4, there is no evidence indicating that zebrafish tmt opsins signal via G_q-coupled pathways that are linked to a change in Ca^{2+} signals.

Since the 420 nm Ca^{2+} response might be caused by the presence of unbound retinal, future experiments can be tested with a much lower concentration of the retinal chromophore. In addition, the cells can be washed with phosphate buffer saline (PBS) to remove any excess retinal prior to experimentation. It is important to note that not all cellular responses are visible in Ca^{2+} imaging, as tmt opsins may couple to different G protein cascades. Other *in vitro* assays, such as a cyclic adenosine monophosphate (cAMP) ELISA assay or G protein binding assay, would also be useful as screening tests for binding of specific G proteins (e.g. G_i-type G protein) .


(A)
TMTCh2 short isoform
TMTCh2 long isoform
TMTCh6
TMTCh9
TMTCh10
TMTCh14
TMTCh24
Average DF/F
Time (s)
545nm
480nm
545nm
9-cis (n=3)
all-trans (n=3)
NR (n=3)
-ve control, 9-cis (n=1)
-ve control, all-trans (n=1)
-ve control, NR (n=1)
hOPN4, 9-cis (n=2)

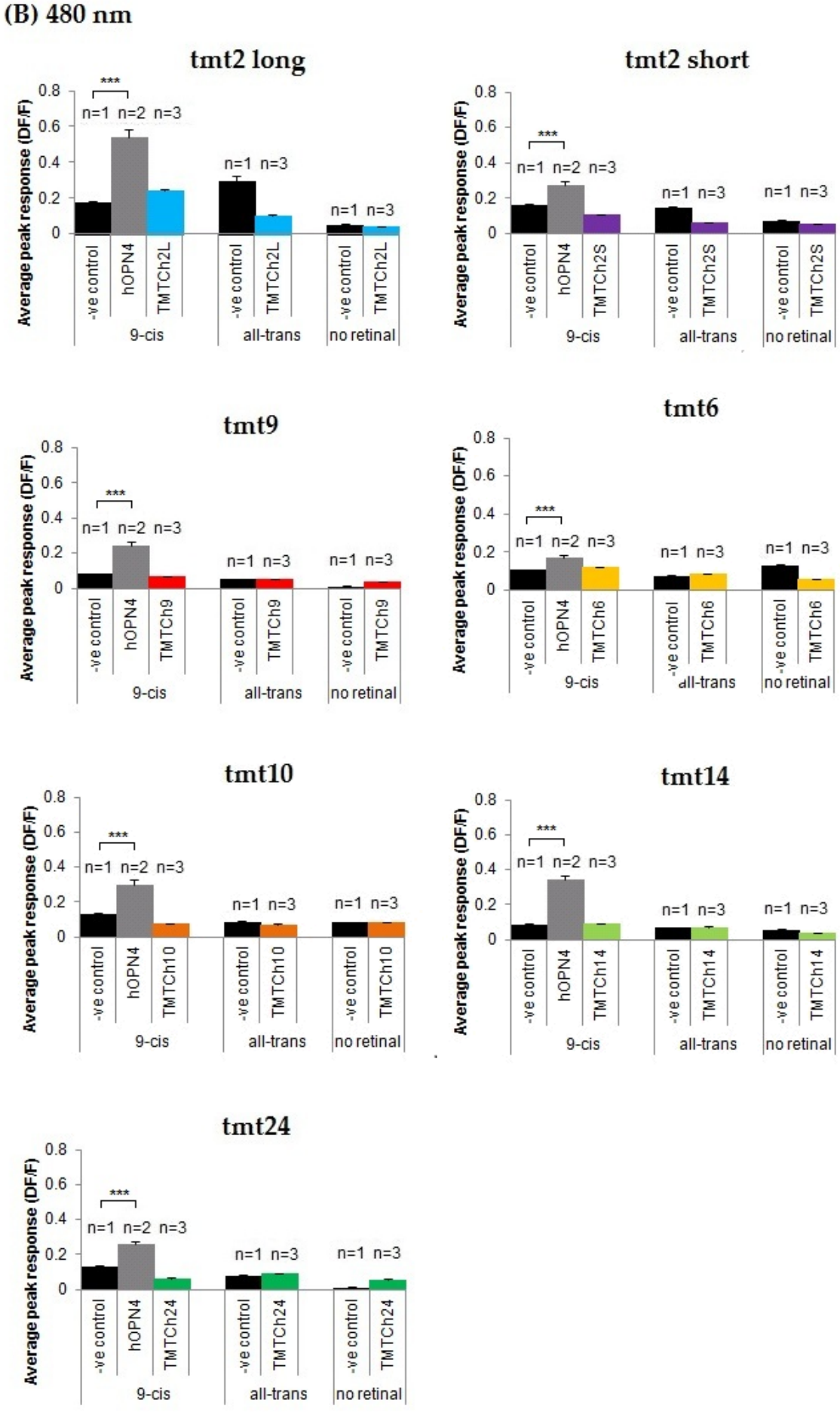
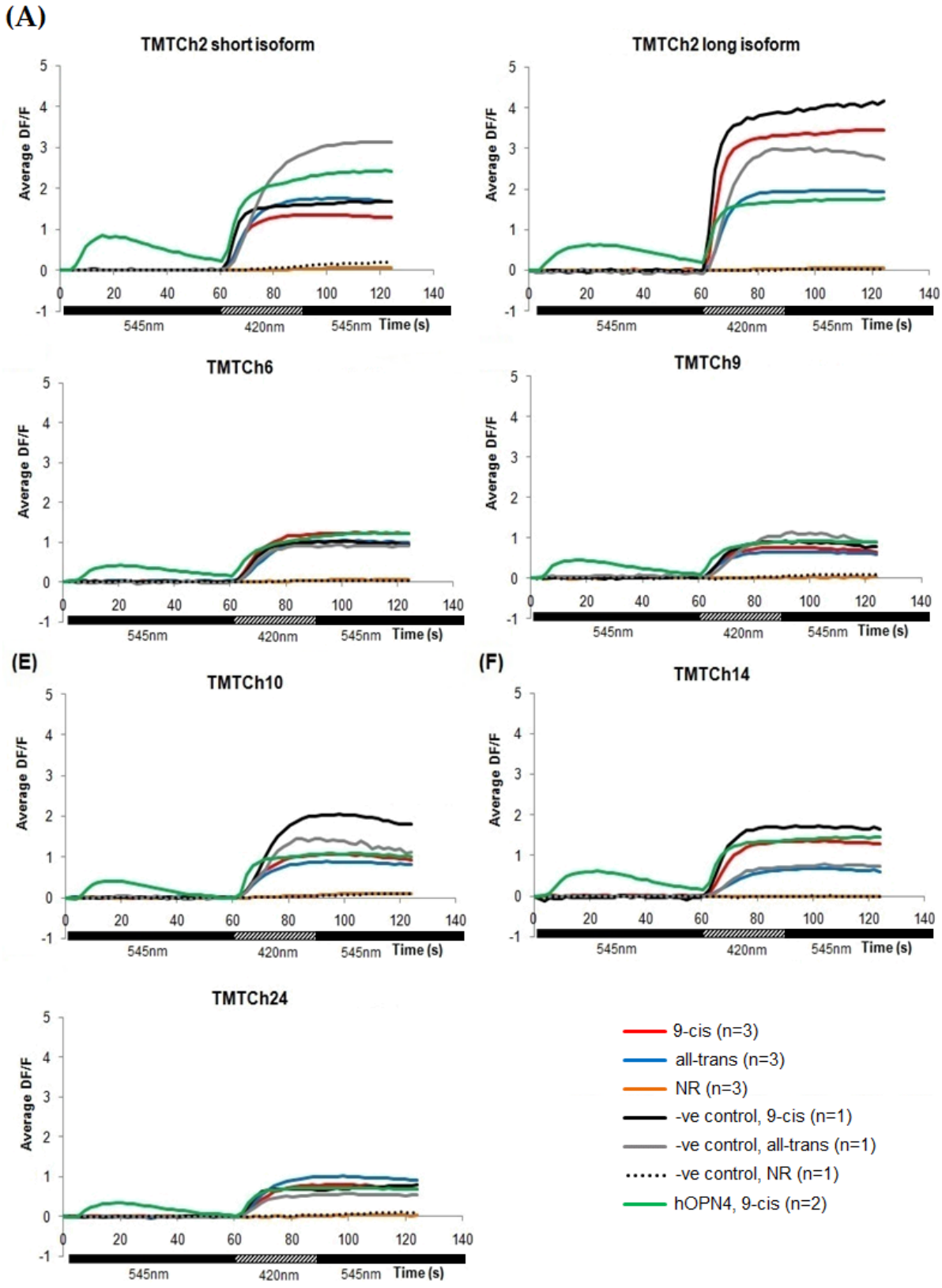

(B) 480 nm
tmt2 long
tmt2 short
tmt9
tmt6
tmt10
tmt14
tmt24
Average peak response (DF/F)
0.8
0.6
0.4
0.2
0
***
n=1 n=2 n=3
n=1 n=3
-ve control
hOPN4
TMTCh2L
TMTCh2S
TMTCh9
TMTCh6
TMTCh10
TMTCh14
TMTCh24
9-cis
all-trans
no retinal

Figure 6.5 (A) Rhod-2 AM calcium imaging of Neuro-2A cells transfected with *tmt2l*, *tmt2s*, *tmt6*, *tmt9*, *tmt10*, *tmt14* and *tmt24*, in the presence and absence of 9-*cis* or all-*trans* retinal. The negative (-ve) controls were untransfected Neuro-2A cells and the positive (+ve) controls were Neuro-2A cells transfected with human OPN4 (hOPN4). Cells were initially pulsed with 100 ms imaging light at 545 nm every 2 s for 30 cycles to check for the presence of the dye, then they were stimulated with 15 x 1 s exposures of 480 nm interspersed between the imaging pulses at 60-90 s, as indicated by the bars below each graph. Traces of representative dF/F measurements for all the cells are shown, averaged from n independent dishes (n indicated in the legend) in response to light stimuli. **(B)** Graphs showing the mean amplitudes of intracellular Ca^{2+} transients in Neuro-2A cells transfected with different tmt opsins upon stimulation by 480 nm (cycles 30-60). Each of the mean amplitudes has a standard error (S.E.) of all the cells measured within a field of view from n independent dishes (n are indicated above the error bars). Statistical analyses were performed using one-way ANOVA with multiple comparisons against untransfected negative control cells by either Dunnett's method for parametric data or Dunn's method for non-parametric data. ***$p < 0.001$


(A)
TMTCh2 short isoform
TMTCh2 long isoform
TMTCh6
TMTCh9
(E)
TMTCh10
(F)
TMTCh14
TMTCh24
Average DF/F
Time (s)
545nm
420nm
545nm
9-cis (n=3)
all-trans (n=3)
NR (n=3)
-ve control, 9-cis (n=1)
-ve control, all-trans (n=1)
-ve control, NR (n=1)
hOPN4, 9-cis (n=2)

(B) 420 nm

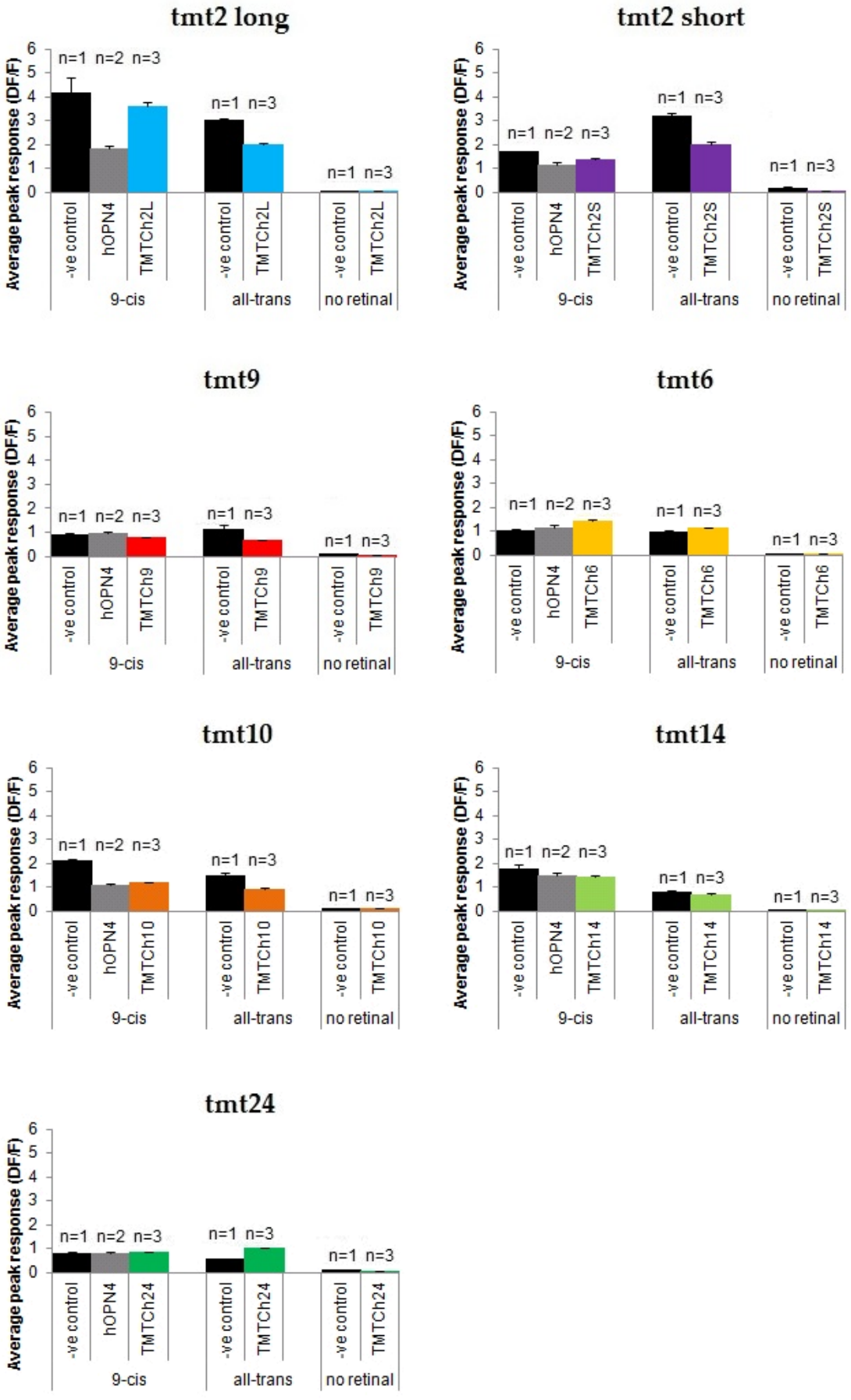

tmt2 long
tmt2 short
tmt9
tmt6
tmt10
tmt14
tmt24
Average peak response (DF/F)
n=1 n=2 n=3
n=1 n=3
-ve control
hOPN4
TMTCh2L
TMTCh2S
TMTCh9
TMTCh6
TMTCh10
TMTCh14
TMTCh24
9-cis
all-trans
no retinal

Figure 6.6 (A) Rhod-2 AM calcium imaging of Neuro-2A cells transfected with *tmt2l*, *tmt2s*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, and *tmt24*, in the presence and absence of 9-cis or all-trans retinal. The negative (-ve) controls were untransfected Neuro-2A cells and the positive (+ve) controls were Neuro-2A cells transfected with hOPN4. Light stimulation protocol is the same as that described in Fig. 6.5A, but with light stimulation 15 × 1 s exposures of 420 nm interspersed between 545 nm imaging pulses at 60-90 s, as indicated by the bars below each graph. Shown are traces of representative dF/F measurements for all the cells averaged from n independent dishes (n indicated in the legend) in response to light stimulus. **(B)** Graphs showing the mean amplitudes of intracellular Ca^{2+} transients in Neuro-2A cells transfected with different tmt opsins upon stimulation by 420 nm (cycles 30-60). Each of the mean amplitudes has a standard error (S.E.) of all the cells measured within a field of view from n independent dishes (n are indicated above the error bars). Statistical analyses were performed as described in Fig. 6.5B.

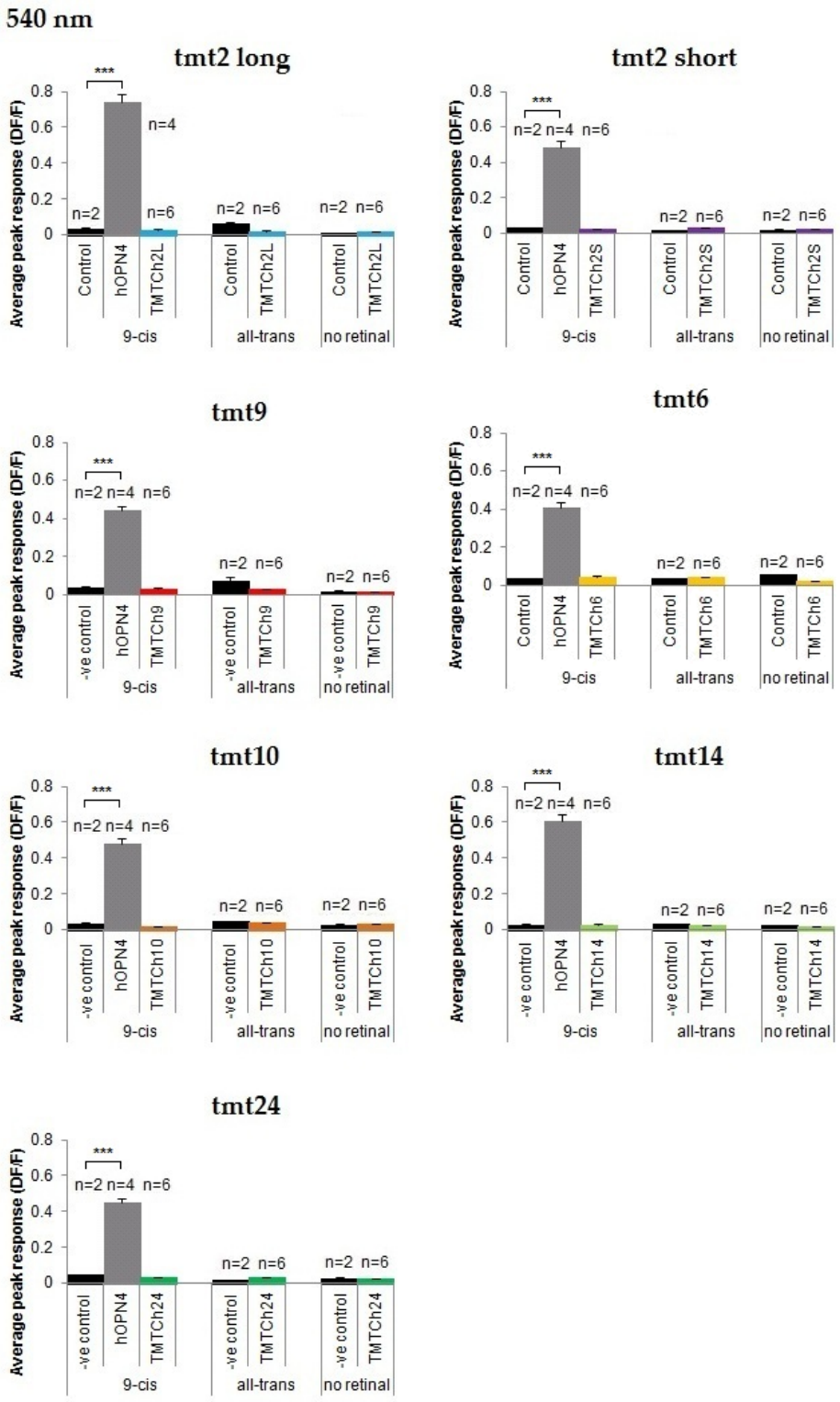


Figure 6.7 Graphs showing the mean amplitudes of intracellular Ca^{2+} transients in Neuro-2A cells transfected with different tmt opsins upon stimulation by 544 nm (cycles 1-30) light pulses. Each of the mean amplitudes has a standard error (S.E.) of all the cells measured within a field of view from n independent dishes (n are indicated above the error bars). Statistical analyses were performed as previously described in Fig. 6.5B.

6.2.3 Electrophysiological studies: tmt-induced photoresponses in Neuro-2A cells

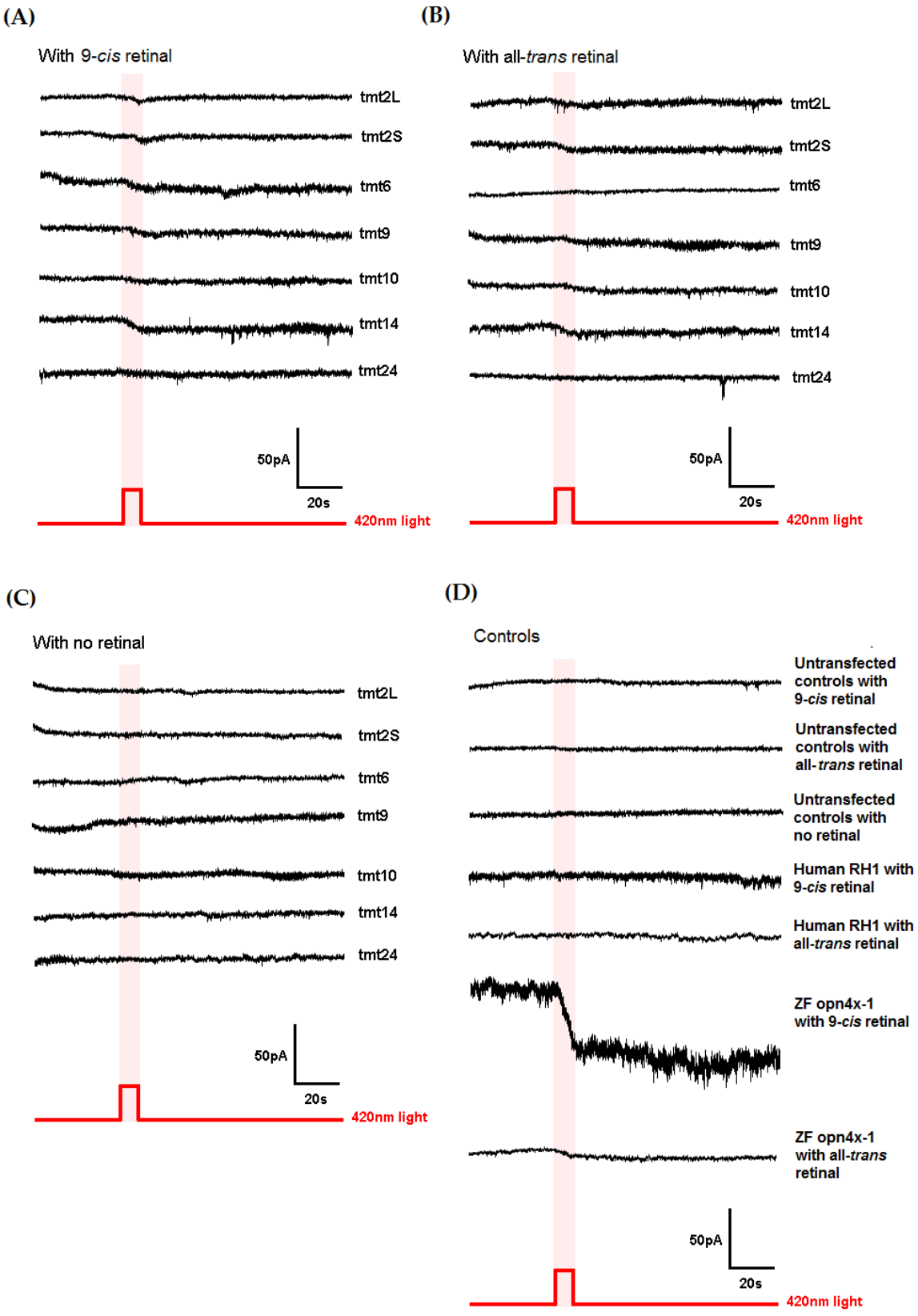
To study the effect of light on membrane potentials in Neuro-2A cells expressing tmt opsins, whole-cell patch clamp recordings were carried out as described in Section 2.5.4.3. Light stimulation at 420 nm was applied to *tmt2s-*, *tmt2l-*, *tmt6-*, *tmt9-*, *tmt10-*, *tmt14-*, and *tmt24*-transfected cells for 10 s to test for the presence of functional opsins. Shorter wavelength from the UV-region (400 nm) was tested, but its damaging effects on the cell membrane made it disruptive for patching as the cells became very unstable. Therefore, only 420 nm light stimulation was used in this experiment. The transfected cells were either treated with 20 μM of retinal (9-*cis* or all-*trans*) for 1.5 h or left untreated prior to experimentation. Zebrafish opn4x-1 was included as a positive control, because studies have shown that its expression in the presence of 9-*cis* or all-*trans* retinal was sufficient to render Neuro-2A cells photosensitive (Davies et al., 2011). Untransfected Neuro-2A cells and $G_{i/o}$-coupled human rhodopsin (RH1) (Cao et al., 2012) were also included as negative controls.

Figures 6.8A-D show the representative current traces recorded from the patching experiment, and Figure 6.8E shows the average changes in membrane current measured in response to the 10 s light stimuli. With the negative controls of untransfected and hRH1-transfected cells, small inward currents between -3.9 (n=1) and -4.43 ± 0.54 pA (n=3) were detected upon light stimulation when cells were preloaded with 9-*cis* and all-*trans* retinals. These light-evoked currents were about 5-6 times greater than those measured in the absence of retinal ($p<0.001$). They were probably intrinsic background responses due to the presence of endogenous opsins (e.g. Opn3 and Opn5), as shown by microarray data analysis of the Neuro-2A cells

(Steven Hughes and Rachel Butler, Oxford). In contrast, the positive control zebrafish opn4x-1-expressing cells exhibited large sustained inward currents of -40.8 ± 5.74 pA (mean ± S.E.M., $n = 6$) with 9-cis retinal, and -12.0 ± 5.12 pA ($n = 4$) with all-*trans* retinal (Davies et al., 2011). These data suggest a retinal-dependent response to 420 nm light in the positive control cells that is significantly greater than background.

When *tmt* opsins were expressed in Neuro-2A cells, small light-evoked responses were also detected in the presence (but not in the absence) of retinal chromophores. However, apart from cells expressing tmt2s, all these light-induced currents were not significantly different to the background responses observed in the untransfected negative controls (see Fig. 6.8E). Although the *tmt2s*-expressing cells pre-incubated with 9-*cis* retinal elicited a slightly larger light-induced current of -5.92 ± 0.48 pA (n=6) compared to the -4 ± 0.17 pA ($n = 3$) of the untransfected cells ($p < 0.05$), the magnitude of the response was eight times lower than the that observed for the zebrafish opn4x-1 positive control. Previous work have shown that this patching assay is reliable for detecting $G_{q/11}$ type activities, and therefore it would seem that tmt2s may not signal primarily via the $G_{q/11}$ pathway.

The small light-induced currents detected in *tmt*-expressing Neuro-2A cells were smaller than those observed in the G_q-coupled Opn4, indicating that they are probably non-specific background responses. Thus, these data suggest that tmt pigments do not signal through a G_q-type G protein signalling cascade. Nonetheless, it remains a possibility that these pigments may act as photoisomerases catalyzing the conversion of retinals between different isomers.


(A)
With 9-cis retinal
tmt2L
tmt2S
tmt6
tmt9
tmt10
tmt14
tmt24
50pA
20s
420nm light
(B)
With all-trans retinal
tmt2L
tmt2S
tmt6
tmt9
tmt10
tmt14
tmt24
50pA
20s
420nm light
(C)
With no retinal
tmt2L
tmt2S
tmt6
tmt9
tmt10
tmt14
tmt24
50pA
20s
420nm light
(D)
Controls
Untransfected controls with 9-cis retinal
Untransfected controls with all-trans retinal
Untransfected controls with no retinal
Human RH1 with 9-cis retinal
Human RH1 with all-trans retinal
ZF opn4x-1 with 9-cis retinal
ZF opn4x-1 with all-trans retinal
50pA
20s
420nm light

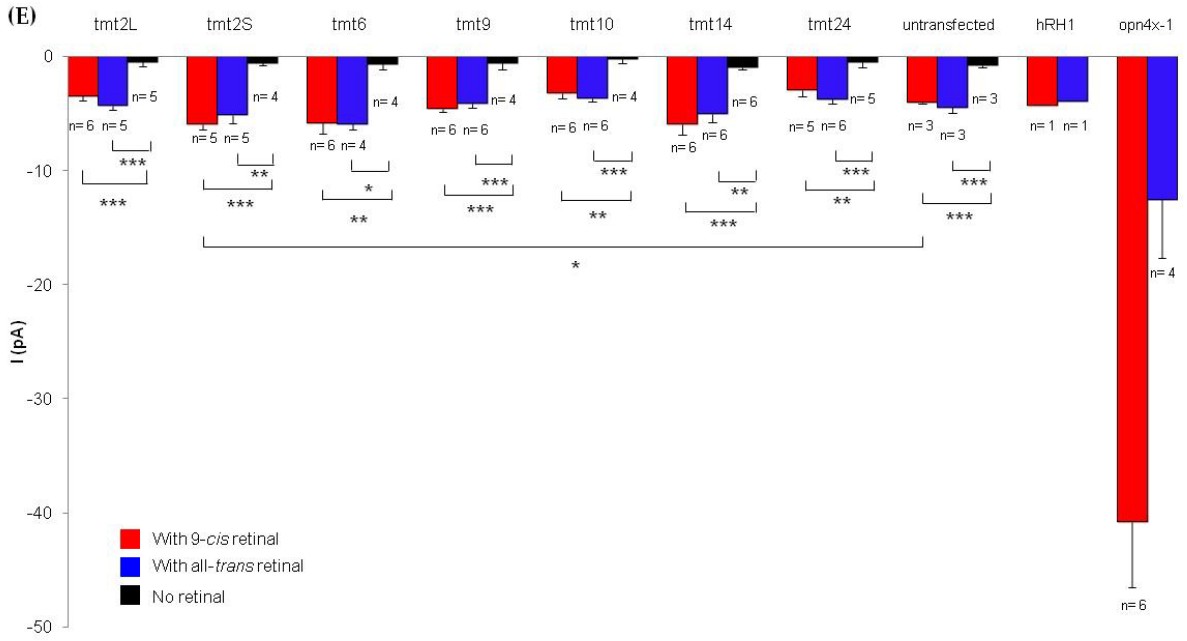


Figure 6.8 Light-evoked responses in Neuro-2A cells expressing six *tmt* opsin coding sequences (*tmt2l*, *tmt2s*, *tmt6*, *tmt9*, *tmt10*, *tmt14* and *tmt24*). **(A-C)** Whole-cell currents recorded from *tmt* opsin-transfected cells pre-incubated with (1) 9-*cis* retinal, (2) all-*trans* retinal, and (3) without retinal. **(D)** Untransfected and human rhodopsin (hRH1)-transfected Neuro-2A cells were included as negative controls. Zebrafish opn4 was used as positive controls. The cells were exposed to 10 s light pulse stimulation at 420 nm (8×10^{14} photons cm^{-2} s^{-1}), which is indicated by the underlying pink regions and the red lines at the bottom of the traces. All cells were clamped at holding potentials of -50mV during recordings. Only representative traces are shown. **(E)** Quantification of peak current amplitudes measured from *tmt* opsin-transfected, untransfected, and control cells upon light pulse stimulation with 9-*cis* retinal (red bars), all-*trans* retinal (blue bars), and no retinal (black bars). Baselines for background light-evoked responses due to the presence of intrinsic opsins are set at -4 ± 0.17 pA for 9-*cis* retinal and -4.43 ± 0.54 pA for all-*trans* retinal based on recordings from the untransfected cells. Data were analysed using student's *t* test for the parametric data, or Mann-Whitney rank sum test if the data were non-parametric. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

6.2.4 G_s, G_i and G_q bioluminescent reporter assays in HEK293 cells

All six zebrafish tmt opsins were subjected to screening by a bioluminescence light response assay to investigate their interaction with G_s, G_i, or G_q-type G proteins (see section 2.5.4.5 for detailed methodology, Bailes et al., 2012). A luciferase-based cAMP reporter (Glo22F; for G_s and G_i assays) and an aequorin-based Ca^{2+} reporter (mt_aq; for G_q assay) were employed in the assay to provide live-cell readout of their respective secondary messenger signals (Fig. 6.9, 6.10, and 6.11A). HEK293 cells were transiently transfected with *tmt2s*, *tmt2l*, *tmt6*, *tmt9*, *tmt10*, *tmt14*, *tmt24* and the reporter plasmids (*Glo22F* or *mt_aq*). The transfected cells were incubated with 9-*cis* retinal (10 μM) overnight prior to the light induction protocol. The light source used in the experiment contained a substantial UV component, encompassing wavelengths from 355 nm to 800 nm (see Fig. 2.4).

For testing of G_s-signal activation, box jellyfish (*Carybdea rastonii*) opsin (JellyOp) was included as a positive control, as it is the only opsin known to signal via the G_s pathway (Koyanagi et al., 2008), whilst untransfected cells were used as negative control. Cells expressing JellyOp showed a 234 ± 18.9-fold increase in cAMP reporter activity induced by a single light flash (355–800 nm, with intensity of 4.5 $mW/m^2/nm$), which rapidly declined to basal level within 5 min (Fig. 6.9). Subsequent repetition of 6 light flashes triggered a bioluminescence response that peaked at 504 ± 46.1 relative light unit (RLU), which was two-fold higher than the initial response measured and took 10 min to return to basal level. Similar cAMP responses had previously been recorded from G_s-coupled JellyOp using the same bioluminescence system (Bailes et al., 2012). In contrast, none of the *tmt*-transfected

cells exhibited a rise in bioluminescence upon light stimulation, suggesting that they do not activate G_s-type G proteins.

As for the G_q-signalling assay, human melanopsin (hOPN4) was used as a positive control. Untransfected cells were also included as negative controls. Cells transfected with hOPN4 displayed a robust Ca^{2+} response to irradiation with a bioluminescence level measured at 295 ± 170.8 RLU (Fig. 6.10A). The kinetics of this bioluminescence response was faster than that of the cAMP experiments, as the Ca^{2+} signal sharply declined to basal level within 1 min. With tmt-expressing cells, none of the tmt opsins tested triggered a Ca^{2+} response to light stimulation (Fig. 6.10A & B) and the bioluminescence levels recorded were similar to those observed in untransfected cells. Therefore, the results suggest that tmt opsins also do not couple to a G_q-type G protein signalling cascade, which is consistent with the Ca^{2+}-imaging and electrophysiological data.

Human rhodopsin (hRH1), a well-characterised G_i-coupled opsin, was used as a positive control in G_i-signalling assay. The presence of hRH1 caused a significant 2-fold decrease in bioluminescence signals with a single light flash when expressed in forskolin-activated HEK293 cells (Fig. 6.11A and B). Subsequent 6 consecutive light flashes caused an even larger inhibition response in hRH1-expressing HEK293 cells, reducing the bioluminescent signal to 5-fold less than that of the untransfected control. Interestingly, some of the tmt-transfected cells responded to light when tested for G_i-binding activity. Light-dependent responses were observed with cells expressing tmt6, tmt10, tmt14, and tmt24, which showed a significant reduction in cAMP reporter signal after one light flash (Fig. 6.11B). However, repeated light

flashes following the single light flash did not enhance these inhibitory activities in tmt-expressing cells (Fig. 6.8C). Hence, the light-dependent G_i-responses mediated by tmt opsins were short-lived and were not immediately reproducible, which could be because of rapid desensitization of the photopigment.

Together, these data demonstrated that within the family of zebrafish tmt opsins, tmt6, tmt10, tmt14, and tmt24, may function as G-protein coupled receptors that specifically activate G_i-type G proteins in a light-dependent manner. Zebrafish tmt opsins do not seem to bind G_s or G_q-type G proteins. These findings correlate with the results obtained from other functional assays, e.g. electrophysiological studies and Ca^{2+} imaging.

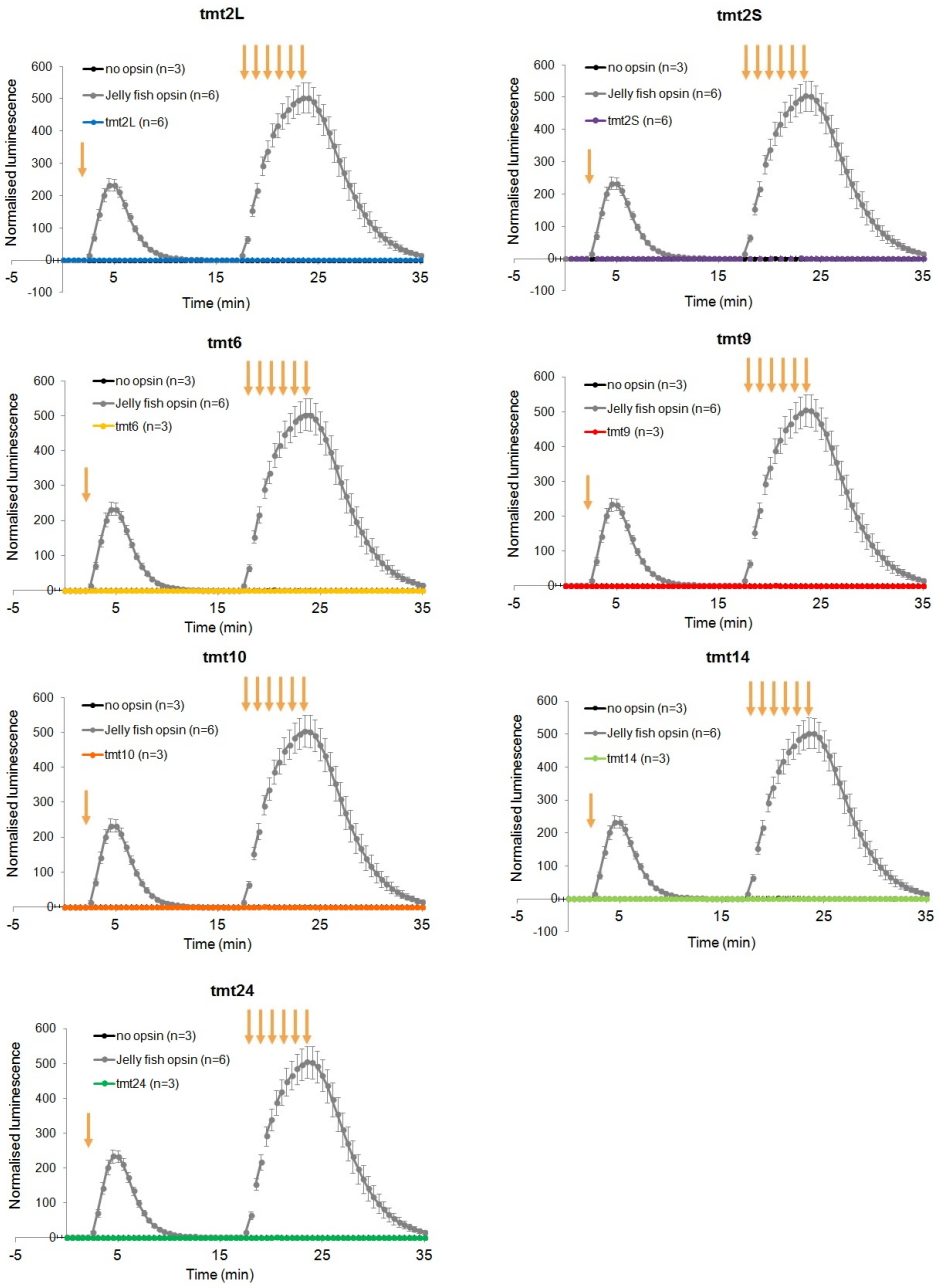


Figure 6.9 Light responses through G_s signalling in HEK293 cells individually expressing each of the six tmt opsins with 9-*cis* retinal. None of the *tmt*-transfected cells exhibited a change in luminescence level upon light stimulation (n=3). Jellyfish opsin (supplied by Dr. Helena Bailes) was used as a positive control, which gave a peak luminescence response (mean ± SEM) of 234 ± 18.9 relative light unit (RLU) after the first light exposure, and 504 ± 46.1 RLU after repeated light exposures (six consecutive flashes, one per minute). Data were recorded and normalised in the same manner as Figure 6.7. The onset of each light flash is depicted by an orange arrow shown above the traces.

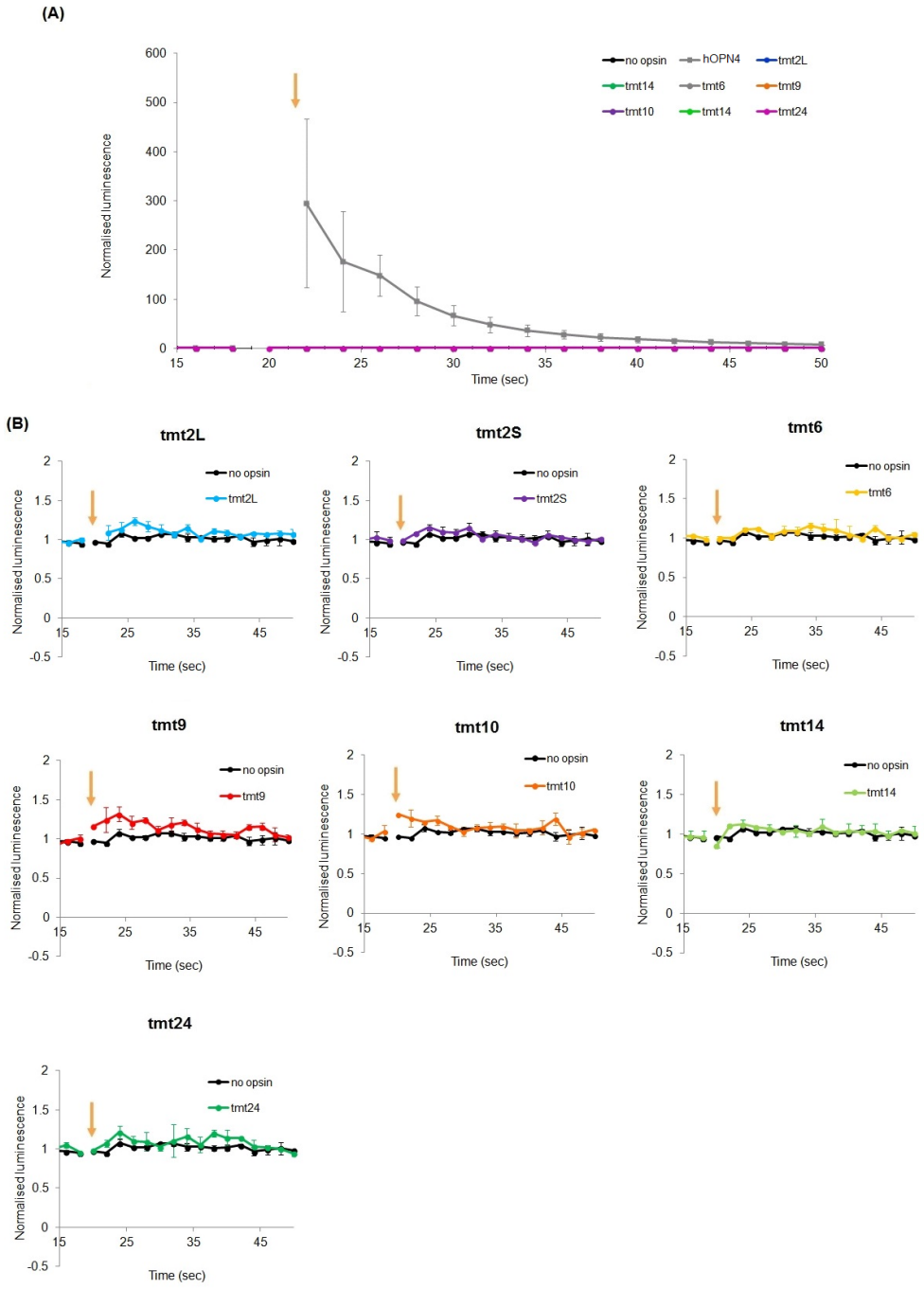
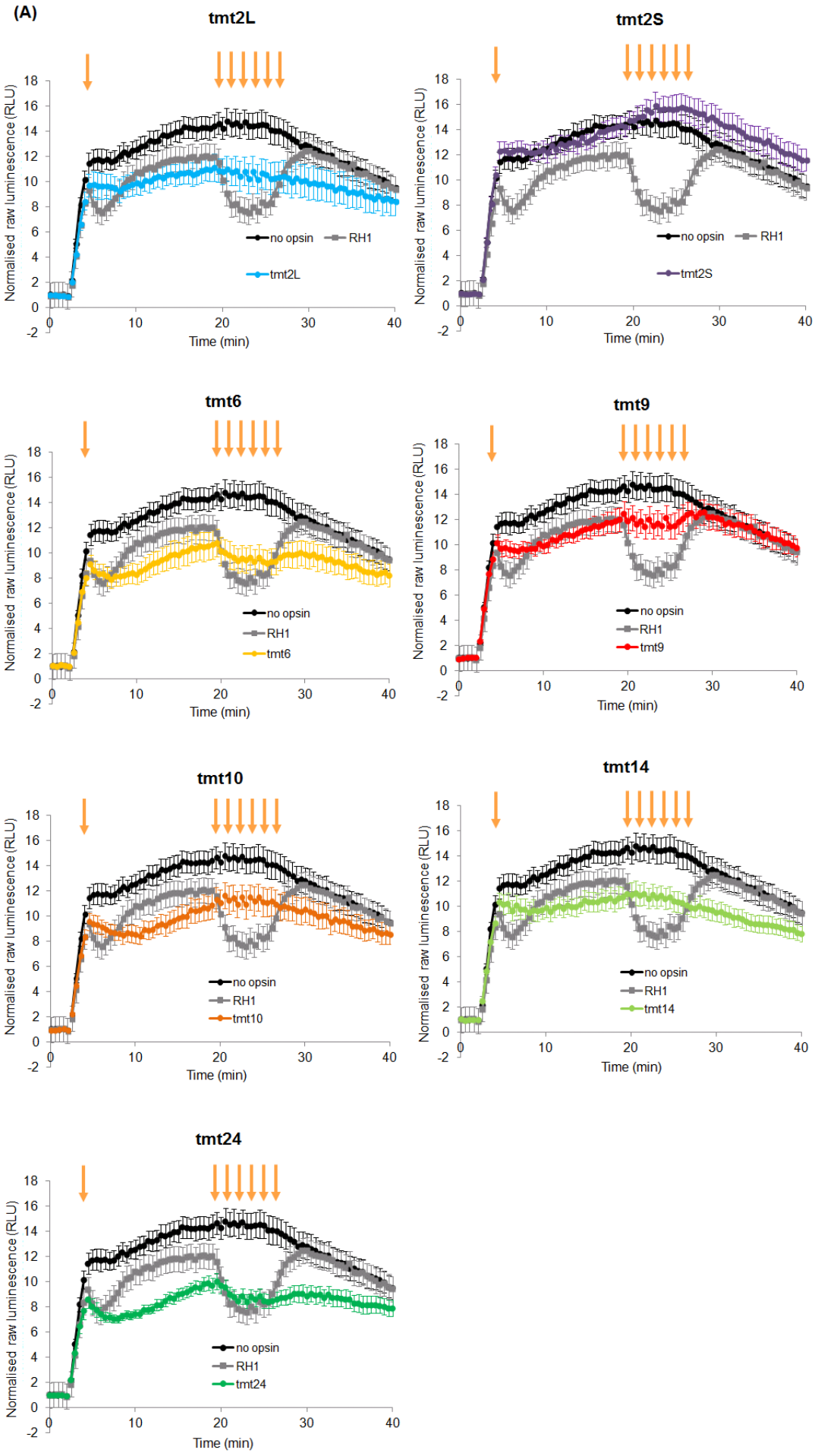


Figure 6.10 Light responses through a G_q-signalling pathway in HEK293 cells for tmt opsins incubated with 9-*cis* retinal. **(A)** Luminescence recordings of a Ca^{2+} biosensor (aequorin) in HEK293 cells expressing the six tmt opsins (n=4). Human melanopsin (hOPN4) was used as positive control and gave a large response measured at 295 ± 170.8 RLU upon light stimulation. Data were recorded and normalised in the same manner as Figure 6.7. Timings of the light flashes are depicted by the orange arrows indicated above the traces. **(B)** Zoomed-in version of the graphs showing the data of *tmt*-transfected cells compared to untransfected controls.


(A)
tmt2L
tmt2S
tmt6
tmt9
tmt10
tmt14
tmt24
Normalised raw luminescence (RLU)
Time (min)
no opsin
RH1
tmt2L
tmt2S
tmt6
tmt9
tmt10
tmt14
tmt24

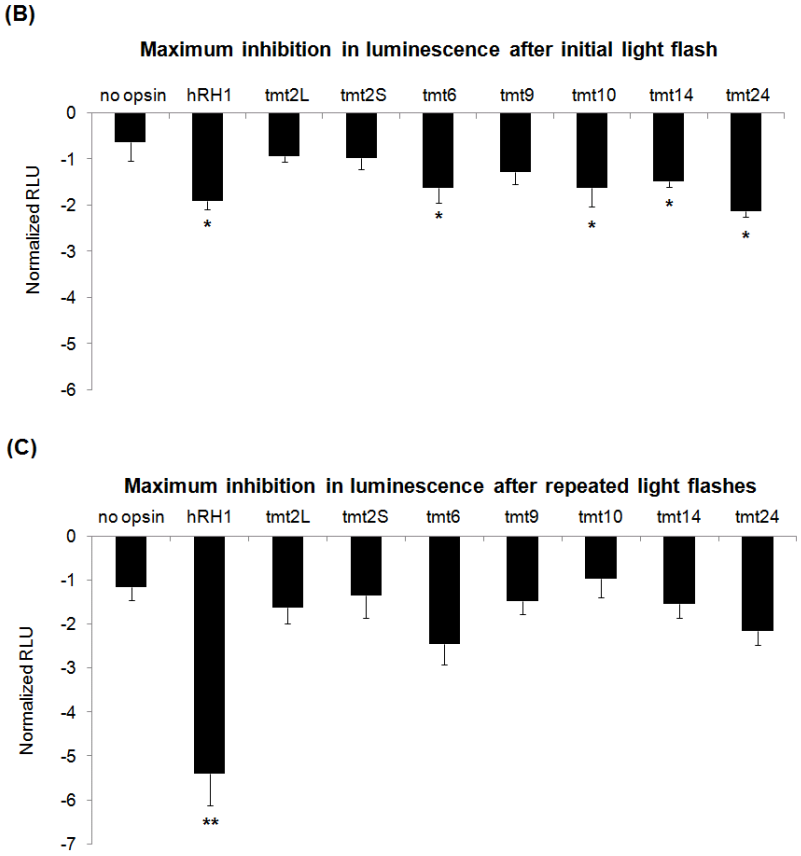


Figure 6.11 Light responses through a G_i-signalling cascade in HEK293 cells driven by tmt opsins with 9-*cis* retinal. **(A)** Light-evoked luminescence changes measured by a cAMP biosensor in HEK293 cells expressing *tmt2l*, *tmt2s*, *tmt6*, *tmt9*, *tmt10*, *tmt14* and *tmt24* (n=9). Data were recorded for 1 s every minute and normalised to baseline levels before light exposure. Timings of the light flashes are depicted by the orange arrows indicated above the traces. **(B)** Quantification of light-induced cAMP inhibitory responses with peak decrease in luminescence after the initial light flash. Human rhodopsin (hRH1) was used as a positive control. The cAMP inhibitory responses were significantly greater in cells expressing *tmt6*, *tmt10*, *tmt14* and *tmt24* than in the untransfected cells. **(C)** Quantification of the sustained responses measured after repeated light exposure (six consecutive flashes, one per minute). The inhibition of cAMP signal was further enhanced in hRH1-transfected cells, but not in the tmt-expressing cells. Statistical analysis in (B) and (C) were performed using one-way ANOVA with multiple comparisons against untransfected cells by either Dunnett's method for parametric data or Dunn's method for non-parametric data. *$p < 0.05$, **$p < 0.01$.

6.3 Discussion

From Ca^{2+}-imaging and electrophysiological studies, there is no evidence for significant activation of G_q signalling cascades by zebrafish tmt opsins (as summarized in table 6.1 below). However, bioluminescence light responses showed that G_i (but not G_q nor G_s) activities were detected with some of the tmt opsins, including tmt6, tmt10, tmt14, and tmt24. Hence, these four tmt opsins have the ability to activate a G_i-type G protein in HEK293 cells when stimulated by light (containing UV wavelengths), which reduce the level of $[cAMP]_i$ possibly by the inhibition of an adenylate cyclase enzyme (Taussig et al., 1993). The bioluminescence reporter assay can be followed up by testing with specific light intensities or wavelengths, e.g. at 370 nm.

Table 6.1 *Summary of the data collected from different functional assays.*

	Ca^{2+} imaging	**Whole-cell patch clamp recording**	**Bioluminescence reporter assays**
Signalling pathway tested	**G_q**	**G_q**	**G_s, G_i and G_q**
tmt2s	×	×	×
tmt2l	×	×	×
tmt6	×	×	G_i-coupled
tmt9	×	×	×
tmt10	×	×	G_i-coupled
tmt14	×	×	G_i-coupled
tmt24	×	×	G_i-coupled

The intracellular surface of an opsin consists of three cytoplasmic loops (CL1, CL2 and CL3) and a C-terminal tail (refer to Section 1.3.3 and Fig. 1.6 in Chapter 1). These cytoplasmic domains possess specific amino acid residues that are involved with the binding of particular G-proteins, opsin kinases, opsin phosphatases and arrestins (Yamashita et al., 2000, Bentrop et al., 1997, Scheerer et al., 2008). They are important for regulating maturation of the opsin protein (Bentrop et al., 1997), selectivity in G-protein coupling and activation of phototransduction cascade (Yamashita et al., 2000). Sequence comparison between the intracellular domains of zebrafish tmt opsins with those of G_i-coupled zebrafish opn5, zebrafish rod opsin (Rh1) and *Drosophila melanogaster* (*D. melanogaster*) Rh1 are shown in table 6.2 below. The analysis revealed that apart from opn5, most of the sequences examined has a stretch of conserved amino acids (LXXPXN) in CL1, two of which are within the sequence (L81 and N86) have been shown to be essential for Rh1 synthesis in *D. melanogaster* (Bentrop et al., 1997). Bentrop and colleagues (1997) showed that mutating these two amino acid residues in *Drosophila* Rh1 disrupted chromophore binding and deglycosylation of the opsin, which produced unstable Rh1 proteins that were subjected to degradation. Although no O-, N-, or C-Linked Glycosylation sites were predicted in the CL1 domain of zebrafish tmt opsins and *Drosophila* Rh1 (using EnsembleGly program, http://turing.cs.iastate.edu/EnsembleGly/predict.html), the presence of these conserved amino acids in tmt opsins may still indicate that these opsins would undergo appropriate post-translational processing for membrane expression. The first seven amino acids at the N-terminal of CL2 in tmt opsins showed a high degree of conservation. Previous studies have demonstrated that

chimaeric mutants of bovine RH1 carrying the CL2 domain of the G_q-coupled muscarinic acetylcholine receptor (mAChR1) completely lacked the wild-type ability in activating G_t and G_o cascades (Yamashita et al., 2000). However, when only the first seven amino acids of this loop in bovine RH1 was replaced, 50% of its ability to activate G_t and G_o was retained. This, therefore, suggests that this region of CL2 is critical for activating G protein cascades (Yamashita et al., 2000). In fact, some conservation were observed in several locations across the CL2 domain of tmt opsins in a class-dependent manner, e.g. at site 136, 142, 147, 149, and 150 (amino acid site numbering based on tmt6). Residues at these sites are highlighted in table 6.2: blue for class I, yellow for class II, and green for class III. Although these findings suggest that the mechanism of activation might be conserved within each *tmt* classes, there is a lack of correlation between G_i-coupling and class-dependency. CL3 is thought to be important for intermolecular recognition between the opsin and specific subtypes of G-protein (Franke et al., 1992, Yamashita et al., 2000). Mutational studies using G_t-coupled bovine RH1 as a template have shown that replacement of CL3 with that found in G_o-coupled receptors, e.g. mAChR1 and prostaglandin E_2 receptor subtype (EP3) can confer G_o activity in the mutant hybrid opsins (Yamashita et al., 2000). A sequence alignment of the CL3 region indicated that indeed, tmt opsins share some similarities with opn5 (e.g. V231 and E245) (numbering based on opn5). In addition, class-specific conservation was also observed in several locations, e.g. site 225, 229, 231, 236, 240, and 241 (numbering based on tmt6). However, there is no specific pattern observed that differentiate the non G_i-coupled tmt opsins (tmt2 and tmt9) from the G_i-coupled ones (tmt6, tmt10, tmt14, and tmt24). The C-terminal tail of each

tmt opsin contains putative palmitoylation and phosphorylation sites that may be involved with post-translational modification of the rhodopsin proteins (refer to Fig. 3.3B). These post-translational processes might serve to enhance the stability of the opsin proteins, influence their membrane localisation (Heck et al., 2003), and/or facilitate their interactions with other proteins such as arrestin (Vishnivetskiy et al., 2007). The extent of conservation within the C-terminus differs significantly across different opsins as well as within each tmt classes, which is not surprising as this region is not implicated to be involved with signalling cascades. Taken together, it may be postulated that some of the tmt opsins which are coupled to G_i-type G protein might employ a new mechanism in activating an endogenous phototransduction cascade.

Table 6.2 *Sequence alignments of cytoplasmic loops and C-terminus of D. rerio (Dr) tmt opsins, opn5, Rh1, and D. melanogaster (Dm) Rh1.*

Cytoplasmic domains	G-protein	Opsins	Sequence alignment	% similarity to opn5
Loop 1	G_i	Dr opn5	TFKRKTKLKPPE	100
	G_i	Dr Rh1	TIEHKKLRTPLN	25
	G_i	DM Rh1	FATTKSLRTPAN	16
	G_i	Dr tmt6	FCKFKTLRTPVN	33
	?	Dr tmt9	FCKFKTLRTPVN	33
	G_i	Dr tmt10	FARFHVLRTPIN	8
	G_i	Dr tmt14	FARFRTLWTPIN	17
	G_i	Dr tmt24	FCRYKVLRSPMN	17
	?	Dr tmt2l	FGRYKVLRSPIN	17
	?	Dr tmt2s	FGRYKVLRSPIN	17
Loop 2	G_i	Dr opn5	DRYLKICHLRYGTWLKRHH	100
	G_i	Dr Rh1	ERWMVVCKPVSNFRFGENH	16
	G_i	DM Rh1	DRYQVIVKGMAGRPMTIPL	21
	G_i	Dr tmt6	DRYSTLTVYNKR~APDYSK	17
	?	Dr tmt9	DRYSTLTVYHKR~APDYRK	17
	G_i	Dr tmt10	ERYSTILCSSKADASDYRK	16
	G_i	Dr tmt14	ERYAALLRATKADVSDFRR	11
	G_i	Dr tmt24	ERYCTMMGSTQADSTNYRK	11
	?	Dr tmt2l	ERYCTMMGSTEADATNYKK	11
	?	Dr tmt2s	ERYCTMMGSTEADATNYKK	11
Loop 3	G_i	Dr opn5	VMIIF~~~~~~~~~KVKSSAKEVSHFDTRNKNNHSLEM	100
	G_i	Dr Rh1	G~~~~~~~~~~~~~RLVCTVKEAAAQQQESETTQRAER	10
	G_i	DM Rh1	WFIIAAVSAHEKAMREQAKKMNVKSLRSSEDAEKSAEG	10
	G_i	Dr tmt6	S~~~~~~~~~~~~~RLICAVKQVGRIR~~KTAARRREY	31
	?	Dr tmt9	G~~~~~~~~~~~~~RLLYAVKQVGKIR~~KTAARKREY	13
	G_i	Dr tmt10	G~~~~~~~~~~~~~KILIAIHGVAKIN~~QTAAQRRET	13
	G_i	Dr tmt14	G~~~~~~~~~~~~~KILLLIKGVTKIN~~LLTAQRREN	17
	G_i	Dr tmt24	G~~~~~~~~~~~~~RLLQAITQVSRIN~~TVVSRKREQ	13
	?	Dr tmt2l	G~~~~~~~~~~~~~KLLHAIKQVSSVN~~TSVSRKREH	22
	?	Dr tmt2s	G~~~~~~~~~~~~~KLLHAIKQVSSVN~~TSVSRKREH	22
C-terminus	G_i	Dr opn5	QVIDCKKKCVKSCCFQAWRKKKPSKTSRFYTISGSIKQRPADEASIEI*	100
	G_i	Dr Rh1	ICMN~~~~~~~~KQFRHCMITTLCCGKNPFEEEEGASTTASKTEASSVSSSSVSPA*	0
	G_i	DM Rh1	GISHPKYRLALKEKCPCCVFGKVDDGKSSDAQSQATASEAESKA*	0
	G_i	Dr tmt6	ILMN~~~KQFYRCFLILIHCKHSSLENGQSSMPSRTTGIQLNRRPYSNPVADNAP~~~~~~~~~~~~~~~~~~~~PSIDLQN~DCSTPVSG*	6
	?	Dr tmt9	ILMN~~~KQFYRCFRILFCCQRSLLQNGHSSMPSKTTVIQLNRRVNSNAVACTAQ~~~~~~~~~~~~~~~~~~~~ISTGTHNHDCSTHVTERSNPPEVIP*	3
	G_i	Dr tmt10	VLFN~~~NQFYRCFIALVRSGAEPPVH~~~~LTLHTEEGAAQQHCIPMGLYAATS~~~~~~~~~~~~~~~~~~~~PPE~~~~~SPSLMDTPK*	7
	G_i	Dr tmt14	VLFN~~~NQFYRCFVAFLKCQGEPSVHGQNPQHSSKEDPHVFRPCDGASIHRSAE~~~~~~~~~~~~~~~~~~~~GPQKKEQHSLSLVVHYTP*	2
	G_i	Dr tmt24	IFMN~~~KQFCRCFHALIMCTTPERGSSFKNSSKVTKTLRTVRRANGQNVTFAVASAVHRTPYSDRQKSSSEGEKLPPATGQGTSKPVVSLVAYYNG*	9
	?	Dr tmt2l	IFMN~~~KQFYRCFRALLNCDKPQRGSSLKSSSKT~KPFRPGRRTD~~NFTFMVASVGP~~~~~~~~NQTNPVEDGPPSADN~TKPAVLSLVAHYNG*	13
	?	Dr tmt2s	IFMN~~~KQVS~MWKDIAGF*	0

Table 6.2 Sequence comparison of the intracellular domains of *D. rerio* tmt opsins with *D. rerio* opn5 (GenBank Accession Number: AY493740), *D. rerio* Rh1 (AB087811) and *D. melanogaster* RH1 (NM079683). The opsin sequences were aligned using ClustalW (Higgins et al., 1996). Conserved domains across all the examined sequences are highlighted in grey, and residues that appear to be conserved in class-dependent manner are highlighted in colours (Class I- blue, Class II- yellow, and Class III- green).

In zebrafish, tmt6, tmt10, tmt14, and tm24 are present in the eye, brain, testis, gut, gill and fin (Fig. 4.2B). They are specifically localised to ocular bipolar cells, amacrine cells and retinal ganglion cells (Fig. 4.5A & B), as well as to various nuclei across different regions of the brain from the telencephalon to the medulla oblongata (table 4.1). These data therefore suggest that zebrafish tmt6, tmt10, tmt14 and tmt24 might function as UVS photopigments that utilise a G_i signalling cascade in these tissues. The other two tmt opsins, tmt2 (from tmtIII) and tmt9 (from tmtI), have also been shown to be UVS (see Chapter 5). However, the present functional data indicated that they may not couple to the G-protein signalling pathways investigated in this study. It is possible that tmt2s, tmt2l and tmt9 may be constitutively active, activate other G-protein (e.g. G_o, G_{12}, G_v), or they may have other non-signalling roles such as acting as photoisomerases to mediate the regeneration and trafficking of retinal in response to light (Bailey and Cassone, 2004).

As mentioned in chapter 1, SWS1 cone opsin is a visual opsin that is also UVS in teleost fish. Reconstitution experiments and G-protein activation assay have demonstrated that this opsin activates the transducin (G_t) cascade (Starace and Knox, 1997, Kusnetzow et al., 2004, Imamoto et al., 2013), therefore indicating a different biological function to that of the G_i-coupled tmt opsins. Parapinopsin (PP-opsin) and opn5 are the only vertebrate non-visual opsins currently reported to have UV sensitivity (Kojima et al., 2011, Yamashita et al., 2010, Wada et al., 2012). Both of these opsins have recently been shown to activate G_i following UV irradiation (Kojima et al., 2011, Yamashita et al., 2010), similar to the findings of zebrafish tmt opsins here. However, neither PP-opsin nor opn5 has an expression pattern in teleosts as broad as

that of the tmt opsins, which encompasses the majority of the tissues in zebrafish (Chapter 4), although many studies have not investigated multiple tissues in such a manner as reported in this study. Currently, PP-opsin expression is confined to the pineal gland and parapineal organ of lampreys and lizards (e.g. the iguana) (Koyanagi et al., 2004, Kawano-Yamashita et al., 2007, Wada et al., 2012), whereas Opn5 is localised to the retina and brain of chicken and mouse, and in the outer regions of the ears of mice (Kojima et al., 2011, Yamashita et al., 2010). All these regions, including the pineal, the eye, and the peripheral tissues of mice (e.g. lung, liver, and skeletal muscle) (Yamazaki et al., 2000) and zebrafish (Whitmore et al., 1998), have been shown to exhibit circadian rhythmicity. Since there is a strong correlation between the expression of circadian clocks and the localisation patterns of these opsins (Yamashita et al., 2010), it suggests that PP-opsin and opn5 may play a role in UV-dependent entrainment of circadian rhythmicity. This study, therefore, expands this hypothesis and suggests that tmt may also play a role in photoentrainment in teleosts.

Studies have shown that the secondary messenger system involving cAMP plays a crucial role in the regulation of circadian clock, where the rhythmic transcription of clock genes is mediated by the cAMP-response element binding protein (CREB), a transcription factor which activates and binds a cAMP-response element (CRE) in the presence of the secondary messenger (Burnside and Ackland, 1984, Wang and Zhou, 2010). The daily oscillation of cAMP signals is therefore closely interlinked with the rhythmicity of circadian clocks (as illustrated in Fig. 6.12). Even though CRE motif (TGACGTCA, Montminy et al., 1986) was not found in

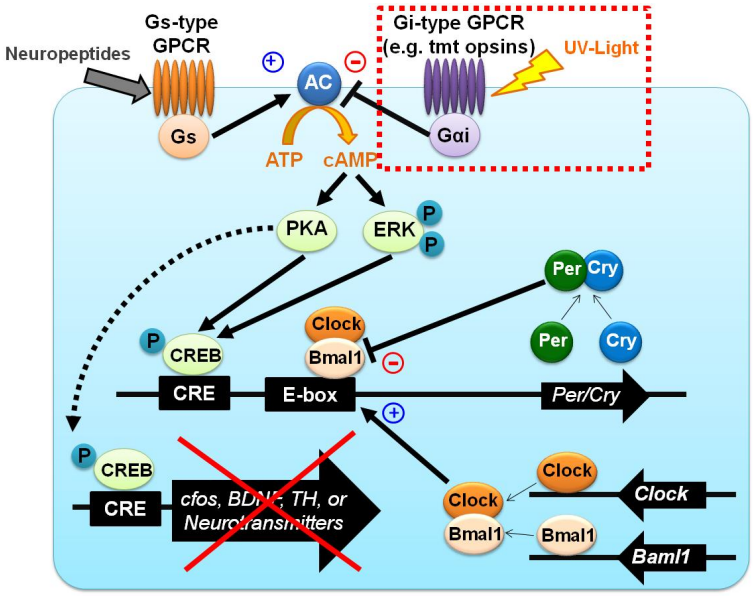


Figure 6.12 A schematic diagram showing both G_s and G_i signalling pathways, and the potential effector genes associated with them. Neuronal clock gene expression can be linked to the rhythmic activation of the cyclic AMP response element (CRE), which is controlled by $[cAMP]_i$. The oscillation of the $[cAMP]_i$ level is in turn regulated by the synchronised activities of both G_s and G_i signalling pathways. Activation of a G_s signalling pathway stimulates adenylate cyclase (AC) to produce cAMP from adenosine triphosphate (ATP), whilst activation of a G_i signalling pathway inhibits AC and decreases $[cAMP]_i$. The balancing of the two signalling cascades in response to external cues (e.g. light-dark cycle, hormones) is essential for the maintenance of clock rhythmicity and synchronisation (Aton et al., 2006). Additionally, binding of CREB to CRE can also transactivate a variety of other genes (e.g. FBJ murine osteosarcoma viral oncogene homolog (*cfos*), brain-derived neurotrophic factor (*BDNF*), tyrosine hydroxylase (*TH*)) that have biological roles in cell growth, differentiation, and protein biosynthesis. Inhibition of the cAMP/CRE system by G_i activation in response to UV-light (highlighted by a red box) might, therefore, be a protective mechanism to enhance cell survival via growth-arrest and cellular senescence. AC, adenylate cyclase; CRE, cAMP responsive element; CREB, CRE-binding protein; P, phosphate; ERK, extracellular signal-regulated kinase; PKA, cAMP-dependent protein kinase; *Per*, period gene; *Cry*, cryptochrome gene; Per, period protein; Cry; cryptochrome protein.

the putative promoters of tmt2 and tmt6, they may be present in the promoter region of other tmt opsins that have not yet been investigated. It is evident that the activation of G_s signals is driven by the circadian release of peptide hormones such as serotonin in the eye (Eskin et al., 1982) or vasoactive intestinal polypeptide (VIP) in the brain (Aton et al., 2006), whilst activation of G_i signals in teleosts could potentially be mediated by photopigments such as tmt opsins, PP-opsin or opn5. G_s activity stimulates AC and cAMP production, whereas G_i signalling inhibits AC and reduces $[cAMP]_i$. Although the activation of G_s and G_i confer opposing effects on the concentration of intracellular cAMP, both pathways are involved simultaneously in controlling rhythmic $[cAMP]_i$ expression (Doi et al., 2011). Blocking either of them pharmacologically abolishes rhythmicity in neurons and desynchronises the system (Aton et al., 2006, O'Neill et al., 2008). The cAMP signal and its oscillation pattern are therefore bona fide components of the circadian clockwork.

Significantly, numerous research studies have demonstrated that there are many other genes whose transcription activities are also controlled by cAMP signals and CRE (Fig. 6.12). These include: FBJ murine osteosarcoma viral oncogene homolog (*c-fos*), a proto-oncogene that has been implicated in cell proliferation, differentiation and transformation (Udagawa et al., 1996, Zhang et al., 2002, Manna and Stocco, 2007); the brain-derived neurotrophic factor (*BDNF*) gene, which is important for neuronal survival, maturation and plasticity (Conti et al., 2002, Tabuchi et al., 2002, Thoenen, 1995); the tyrosine hydroxylase (*TH*) gene, an enzyme that is involved in the biosynthesis of neurotransmitters e.g. dopamine, norepinephrine, and epinephrine (Lewis-Tuffin et al., 2004); and various neuropeptides genes such as

somatostatin (Gonzalez and Montminy, 1989) and VGF nerve growth factor (Di Rocco et al., 1997) genes. Therefore, G_i-activation would serve to reduce the intracellular cAMP signal and suppresses the expression of the genes mentioned above. Apart from potentially being involved in circadian photoentrainment, tmt opsin-mediated G_i-signalling may also enable an organism to detect UV-light for a variety of survival functions, such as DNA repair, growth arrest and protection from cell death in damaging light environments. It can be postulated that tmt opsins may associate with photolyase enzymes (e.g. *phr6-4*; DNA (6-4) photolyase), which function to detect and repair damaged DNA caused by exposure to UV light (Kobayashi et al., 2000). These photolyases have the ability to bind to specific pre-mutagenic lesions induced in the DNA by UV and catalyse the breakage of harmful pyrimidine dimers upon light activation, converting them back to the original bases as a protective mechanism for the cells (Todo et al., 1993). Interestingly, phylogenetic analyses showed that both *tmt* opsin (Section 3.2.3) and DNA photolyase (Kato et al., 1994, Lucas-Lledo and Lynch, 2009, Kuraku and Kuratani, 2011) genes were lost in the eutherian mammals, suggesting the likelihood of functional dependency between these two sets of genes.

There is a possibility that two of the zebrafish tmt opsins in this group, tmt2 (long and short isoforms) and tmt9, may function as retinal-dependent photoisomerases because they do not seem to couple with any of the G-protein pathways tested under these experimental conditions, although there are other G proteins that may be involved. Photoisomerases are proteins that catalyse the regeneration of 11-*cis* retinal from all-*trans* retinal in a light-dependant manner. The

best described photoisomerases are RGR opsin and squid retinochrome. RGR opsin is present in the RPE and Müller glial cells of the vertebrate retina (Jiang et al., 1993), and retinochrome is localised to the inner and outer segments of the mollusc visual cell (Hara and Hara, 1976). RGR and retinochrome absorb light at 520 nm to convert all-*trans* retinal into 11-*cis* configuration (Hara et al., 1967, Shen et al., 1994). Similarly, tmt2 and tmt9 may absorb wavelengths in the UV range to mediate the same photoreaction on retinal chromophore. However, instead of being confined within the retina, this process of retinal regeneration by tmt opsins may occur in a broad range of tissue types including the eye, brain, testis, gut, heart, and fin (see Fig. 4.2). To test for potential photoisomerase characteristics of tmt opsins, changes in retinoid content of primary cultured cells isolated from these tissues following light exposure may be examined. Concomitantly, analysis of the retinoid content extracted from homogenised cells transfected with *tmt* opsins and preloaded with a particular chromophore may be determined by high-performance liquid chromatography (HPLC) subsequent to UV light treatment (Radu et al., 2008). If tmt-mediated photoisomerase activity is present, there should be a lower level of free all-*trans* retinal detected in the UV-exposed cells compared to those that are not stimulated by light, as the all-*trans* retinal would have been converted into 11-*cis* isomer. Nevertheless, it is important to note that tmt opsins may also bind to the 11-*cis* retinal generated and convert it back to all-*trans* form, since they appear to function as bistable pigments (refer to UV-vis spectrophotometric data in Section 5.2.1). Therefore, it is likely that a steady-state mixture of 11-*cis* and all-*trans* retinals would be reached.

All in all, the results of this study suggest that tmt6, tmt10, tmt14 and tmt24 may specifically mediate G_i-type G protein signalling, whereas tmt2 (long and short isoforms) and tmt9 may have other functions that do not involve G protein signalling such as photoisomerisation when triggered by UV light or act as a binding protein for both 11-*cis* and all-*trans* retinal. Some proteins, e.g. interphotoreceptor retinoid-binding protein (IRBP) and ATP-binding cassette transporter gene sub-family A (ABCA4), bind retinoid and are involved in their transport (Okajima et al., 1990, Wolf, 1998, Zhong and Molday, 2010). It could be that tmt2 and tmt9 has the same function in storing excess retinoid which is damaging to the cell. Irradiation with UV light may allow release of chromophore from the opsin-bound stores, and the photochemical reaction does not have to be involved with signalling or photo-conversion of retinal. This may also explain the broad expression profile of tmt opsins. It can be speculated that activation of a G_i pathway may act as a feedback mechanism to antagonise the activity of G_s signalling, maintaining homeostatic processes in the circadian system. Nonetheless, this G_i signalling pathway may also control a diverse range of other cellular functions that might be fundamental to the survival of an organism in a potentially harmful UV-rich light environment.

CHAPTER 7

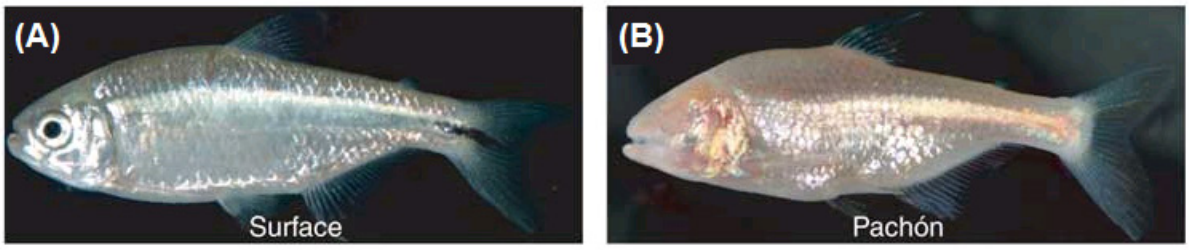
Identification of tmt Opsins in Cavefish

CHAPTER 7: Identification of tmt Opsins in Cavefish

7.1 Introduction

The Mexican tetra, *Astyanax mexicanus*, is a species of fresh water fish that comprises of a sighted epigean form (surface fish, Fig. 7.1A) and over 30 different populations of blind hypogean forms (cavefish, Fig. 7.1B) living in isolated geographic regions of North-Eastern Mexico (Mitchell et al., 1977, Protas et al., 2008). The two cave populations investigated in this study, Pachón and Steinhardt, are from the Pachón and the Chica caves, respectively. These two caves have separated internal drainage systems (Gross, 2012), and their geographic locations are illustrated in Fig. 7.1C. Studies have indicated that the present day populations of northern caves (e.g. Pachón) originated from a surface stock that was distinct, and more ancient to the one which gave rise to the southern cave populations (e.g. Chica) (Trajano et al., 2010).

Remarkably, the hypogean forms of *A. mexicanus* have all evolved similar troglomorphic phenotypes from living in a dark cave environment, with degenerated eyes and a reduction in melanin pigmentation (see Fig. 7.1B) (Trajano et al., 2010). These cavefish initially develop eye primordia, which begin to degenerate 48 hpf, triggered by apoptosis of the lens (Alunni et al., 2007). Interestingly, despite the complete absence of light in deep cave environments and the lack of a functional eye structure, various studies have shown that *A. mexicanus* exhibits photoresponses in their swimming activity (Yoshizawa and Jeffery, 2008, Erckens and Martin, 1982). It has been suggested that these light responses are mediated by photoreceptors in the



(C)

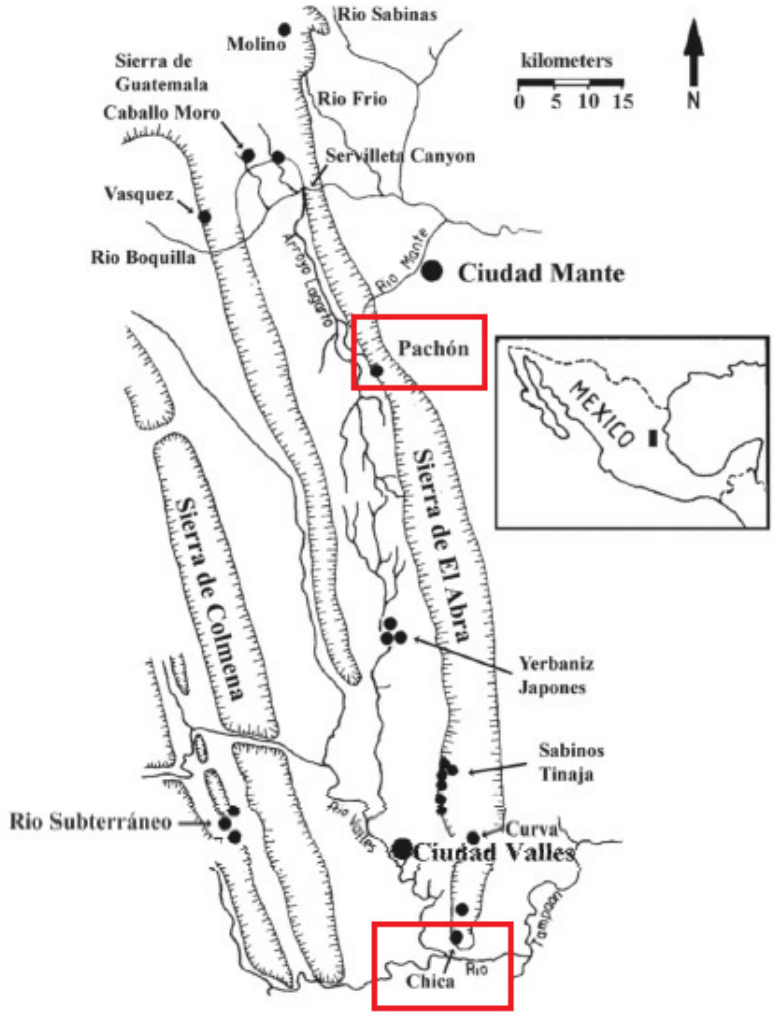


Figure 7.1 Representative photos of *A. mexicanus* from **(A)** the surface population and **(B)** the Pachón cave population (Protas et al., 2008). **(C)** A map of North-Eastern Mexico showing origins of the two cave populations studied, which are highlighted by red boxes (Trajano et al., 2010). The Pachón cave is located about 60 km away from the Chica cave.

pineal complex (Ladich, 2006, Yoshizawa and Jeffery, 2008), although the identity of the photopigment(s) or the photoreceptor(s) has yet to be elucidated. Evidence for intact opsin genes and photoreceptors have been found in the pineal of blind cavefish *A. mexicanus* (Parry et al., 2003). This suggests that despite living in a dark environment, functional opsins may be present in these teleost fish but they act to regulate the pineal instead of the eye.

Like zebrafish, *A. mexicanus* expresses circadian clock genes throughout development (from 0-72 hpf) (Beale et al., unpublished data) and into adulthood, e.g. in the degenerating eyes of developing cavefish (Espinasa and Jeffery, 2006), as well as in the peripheral tissues such as fin (Beale et al., unpublished data). The rhythmicity of *A. mexicanus* clock gene expression oscillates with a period that is similar to that observed in zebrafish (Beale et al., unpublished data). Significantly, the circadian clocks of both zebrafish and cavefish are entrainable by external light cues (Whitmore et al., 2000). Together, these findings suggest the presence of a peripheral photoreceptor that could be responsible for mediating the light responses observed in cavefish behavioural studies and/or possibly the entrainment of their circadian clocks. Compared to zebrafish, a similar peripheral light sensor like tmt opsins may be present in the *A. mexicanus* despite the lack of external light.

Tmt opsin has been proposed as a mediator for the photoentrainment of peripheral clocks in zebrafish, due to its broad expression pattern and conserved molecular features suggestive of a functional photopigment (Moutsaki et al., 2003). Indeed, strong evidence from this project has confirmed that all six members of the zebrafish tmt opsin subfamily are UVS pigments (Chapter 5), with the majority of

them (4 out of 6) maybe coupling and signalling through a G_i protein cascade (Chapter 6). A previous report by Cavallari et al. (2011) suggests that tmt opsin and/or opn4 may provide the circadian input for clock gene expression in *P. andruzzi*. However, only one tmt opsin was examined in that study. The aims of this study were therefore to: (1) elucidate the effect of light on the function and molecular adaptation of all six *tmt* opsins in epigean and hypogean forms of *A. mexicanus*, and (2) find evidence that adds support to tmt opsins being important for peripheral photoentrainment in teleost fish.

7.2 Results

7.2.1. Isolation of tmt opsins in *A. mexicanus* (surface and cave forms)

Three different *A. mexicanus* cell lines were derived from the embryos of surface (SF), Pachón (PC), and Steinhardt (SH) populations as described in Section 2.5.5.1. Using cDNA generated from the mRNA of these cell lines and a set of tmt opsin-specific degenerate primers (Table 2.5), partial sequences of *tmt* opsins were isolated by RT-PCR. The primers used were consisted of three sets of four primers, with each set specific for tmt class I, class II, and class III. A combination of blast searches (blastn and tblastx) and phylogenetic analyses confirmed the identity of these partial sequences as orthologues of zebrafish *tmt9*, *tmt14*, and *tmt24*, and they were subsequently named based on sequence identity of grouping with the specific zebrafish tmt class. These partial sequences and their genomic structures are summarised in Fig. 7.1. *tmt6* was been previously been identified from the surface fish by Andrew Beale (UCL, UK) using the same primers, the sequence was subsequently extended by RACE-PCRs. Full-length sequences of *tmt6* from the other two cave populations were also determined in this study from RT-PCR.

5′ and 3′-RACE PCRs were performed with primers specifically designed for each cavefish partial *tmt* opsin (Table 2.6 and 2.7) in an attempt to obtain full-length sequences. Despite the trial of two different RACE kits (FirstChoice® RLM-RACE Kit- Life Technologies and 5'/3' RACE Kit, 2nd Generation- Roche Applied Science), both 5′-RACE experiments failed to amplify any cDNA of interest. On the other hand, 3′-RACE amplification was successful for deriving the 3′-ends of *tmt9* and *tmt14*. In addition, fragments of a long and a short isoform were identified for *tmt9*,

whereby the short isoform has three exons with the third exon reading into a stop codon in intron 3. Interestingly, isoform variants exist for *tmt9* in *A. mexicanus* but it is *tmt2* that has long and short isoforms in zebrafish. The total sizes of overlapping fragments are: *tmt9l* (898 bp), *tmt9s* (742 bp), *tmt14* (274 bp), and *tmt24* (629 bp). Unlike the *tmt2* isolated from Somalian cavefish, *P. andruzzii* (Cavallari et al., 2011), these *A. mexicanus tmt* opsins do not possess an insertion mutation nor a premature stop codon that would influence its ability to bind retinal chromophore. Sequence details of the partial *tmt9*, *tmt14* and *tmt24*, and the full-length *tmt6* currently obtained are shown in appendices Section A.2.

Genomic long-range PCR was also performed with LongRange PCR kit (QIAGEN) in an attempt to amplify the introns from genomic DNA and to characterize the intron/exon boundaries of *A. mexicanus tmt* opsins. Unfortunately, the genomic PCR also failed to yield specific products from the *A. mexicanus* genomic DNA. The estimated genome size of *D. rerio* is 1.7 billion bases (Kasahara and Morishita, 2006), whilst that of the *A. mexicanus* has not yet been reported. The failure of genomic DNA sequencing in *A. mexicanus* suggests the technique did not work on their genome.

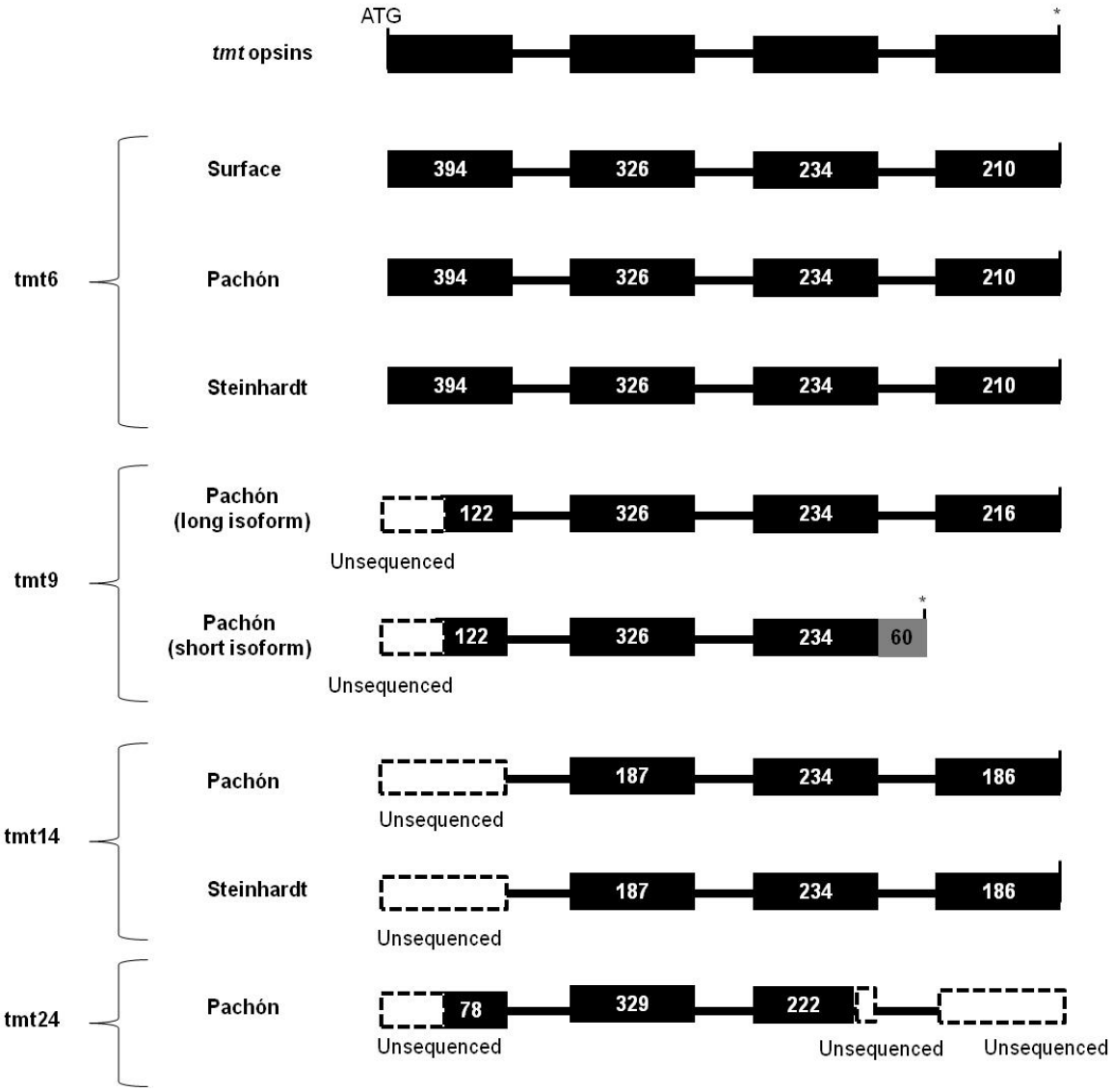


Figure 7.1 Diagram showing the *A. mexicanus tmt* opsin genes predicted based on their orthologues in *D. rerio*. The *A. mexicanus tmt* opsins were isolated from *A. mexicanus* embryonic cell lines derived from surface, Pachón and Steinhardt populations. Exon regions are shown as the black boxes, and regions of undetermined sequences are represented by boxes with dash lined pattern. The size of each exon is marked as base pair (bp) within the boxes. The start (ATG) and stop (*) codons of each genes are indicated in the representative *tmt* opsin in the top row. *tmt6* was the only full-length sequence identified in all three populations. A short splice variant of *tmt9* was identified from its partial sequence, which possesses part of exon 1, complete length of exon 2 and 3, with exon 3 running into intron 3 (grey box) to the next available stop codon located within the intron.

The full-length *tmt6* coding sequences from surface, Pachón, and Steinhardt populations of *A. mexicanus* each contains an open reading frame (ORF) of 1161 bp, which encodes for a protein of 387 amino acids. At both amino acid and nucleotide level, these sequences share over 99% identity amongst themselves and between 70-75% identity to zebrafish *tmt6* (see Table 7.1 below). The small degree of sequence variability amongst the three *A. mexicanus* populations suggests non-adaptive evolutionary changes caused by random mutations and/or genetic drifts. It is not surprising that tmt6 of the two cave forms (Pachón and Steinhardt) are slightly more closely related with each other (99.5% amino acid similarity) than to the surface form (99.2% amino acid identity). Due to the stability of the cave environment, the selection pressure on genetic variations in the hypogean form of *A. mexicanus* is expected to be relatively weaker than that on the epigean form, which is directly exposed to light on the surface water.

Table 7.1 *Nucleotide and amino acid sequence identity shared by the tmt6 of A. mexicanus (surface, Pachón, and Steinhardt), zebrafish tmt6, zebrafish opn3, and bovine rhodopsin.*

		Nucleotide sequence identity (%)					
	tmt opsins	***Bt* Rh1**	***Dr opn3***	***Dr tmt6***	**SF *tmt6***	**PC *tmt6***	**SH *tmt6***
Amino acid identity (%)	***Bos taurus (Bt)* Rhodopsin (Rh1)**	100	37.2	38.3	41.4	41.4	41.4
	Danio rerio* (Dr) *opn3	23.6	100	41.4	44.0	44.1	44.1
	Danio rerio* (Dr) *tmt6	23.9	30.6	100	70.4	70.6	70.6
	Surface (SF) *tmt6*	24.2	32.4	74.7	100	99.5	99.5
	Pachón (PC) *tmt6*	24.2	32.7	74.7	99.2	100	99.5
	Steinhardt (SH) *tmt6*	23.9	32.4	75.0	99.2	99.5	100

7.2.2. Key structural features of *A. mexicanus* tmt opsins

Secondary structures of each full-length (tmt6) and partial (tmt9l, tmt9s, tmt14, and tmt24) *A. mexicanus* tmt opsin sequences were analysed using online software TMHMM Server version 2.0 (see Appendix Section A.4 for results). For the full-length sequences of *A. mexicanus* tmt6, seven transmembrane domains were predicted, as would be expected for a GPCR protein (see Fig. 7.2A). For all other partial tmt opsin sequences of *A. mexicanus*, their transmembrane domains encompass similar regions of amino acids when compared with homologous tmt opsins in zebrafish. These results indicate that *A. mexicanus* tmt opsins are likely to still be structurally intact, and possibly functional (at least over the regions determined) despite the restricted light environment.

Alignment of the deduced amino acid sequences of *A. mexicanus* tmt opsins with zebrafish tmt opsins and bovine RH1 revealed conserved structural and functional features that characterise GPCRs and specifically the opsin family (see Fig. 7.2B). Based on the numbering scheme of bovine RH1 amino acid sequence, these features include: (1) K296; (2) C110 and C187; (3) two putative N-linked glycosylation sites at the N-terminus; (4) putative palmitoylation sites at the C-terminus; (5) E(D)/R/Y motif (position 134-136); (6) NPxxY$(x)_{5,6}$F motif,; and (7) putative phosphorylation sites at C-terminal tail (refer to Section 3.2.2 for specific structural and functional significance). As these features are found across all the opsin classes and are known to play a role in the photosensory function of opsin pigments, it seems that the tmt opsins of *A. mexicanus* retain all the key features of light sensitive opsins even in the cave populations.

(A)

```
Bt_ROD_OPSIN        ---------- -MNGTEGPNF YVPFSNKTGV VRSPFEAPQY YLAEPWQFSM 39
Dr_tmt6             MFPEETNMSY ISNGTD---- ----DDL-LS ALEDWSDTPA EKLSRTGHNV 41
Am_SF_tmt6          MFYDLEISNF STNWSTGEED AEGEEGALSA LGWGWSDAPE QRLTRTGHSV 50
Am_PC_tmt6          MFYDLEISNF STNWSTGEED AEGEEGALSA LGWGWSDAPE QRLTRTGHSV 50
Am_SH_tmt6          MFYDLEIFNF STNWSTGEED AEGEEGALSA LGWGWSDAPE QRLTRTGHSV 50
Dr_tmt9             MFFEQADLNY SFNMSE---- ----EDRLTL LDEDWSDSPM ETLSRAGFIA 42
Am_PC_tmt9L         ---------- ---------- ---------- ---------- ---------- 0
Am_PC_tmt9S         ---------- ---------- ---------- ---------- ---------- 0
Dr_tmt14            ----MVVYIW SLNISSKDTS ALNQSGNVSS GDPLEPHDSP PGLSRTGHTV 46
Am_PC_tmt14         ---------- ---------- ---------- ---------- ---------- 0
Am_SH_tmt14         ---------- ---------- ---------- ---------- ---------- 0
Dr_tmt24            --MIESNVSR SCEWCAGGG- ----EGTGAH LDENHSDH-- -SLSPTGHLV 40
Am_PC_tmt24         ---------- ---------- ---------- ---------- ---------- 0
Clustal Consensus

                                                 First
                                TMI              cytoplasmic loop          TMII
Bt_ROD_OPSIN        LAAYMFLLIM LGFPINFLTL YVTVQHKKLR TPLNYILLNL AVADLFMVFG 89
Dr_tmt6             VAVILGSILI FGTLNNLVVL VLFCKFKTLR TPVNMLLLNI SVSDMLVCLF 91
Am_SF_tmt6          VAVILGFIMV FGFLNNVVVL VLFCRFKTLR TPVNLLLLNI SVSDMLVCVC 100
Am_PC_tmt6          VAVILGFIMV FGFLNNVVVL VLFCRFKTLR TPVNLLLLNI SVSDMLVCVC 100
Am_SH_tmt6          VAVILGFIMV FGFLNNVVVL VLFCRFKTLR TPVNLLLLNI SVSDMLVCVC 100
Dr_tmt9             LSVFLGFIMT FGFFNNLVVL VLFCKFKTLR TPVNMLLLNI SISDMLVCMF 92
Am_PC_tmt9L         ---------- ---------- ---------- ---------- -VSDMLVCTC 9
Am_PC_tmt9S         ---------- ---------- ---------- ---------- -VSDMLVCTC 9
Dr_tmt14            TAVCLGAILL LGCLNNLFVL LVFARFRTLW TPINLILLNI SVSDILVCLF 96
Am_PC_tmt14         ---------- ---------- ---------- ---------- ---------- 0
Am_SH_tmt14         ---------- ---------- ---------- ---------- ---------- 0
Dr_tmt24            VAVCLGFIGT FGFLNNTLVL VLFCRYKVLR SPMNCLLISI SVSDLLVCVL 90
Am_PC_tmt24         ---------- ---------- ---------- ---------- ---------- 0
Clustal Consensus

                                                                  Second
                                                    TMIII         cytoplasmic loop
Bt_ROD_OPSIN        GFTTTLYTSL HGYFVFGPTG CNLEGFFATL GGEIALWSLV VLAIERYVVV 139
Dr_tmt6             GTTLSFAASI RGRWLVGRHG CMWYGFVNSC FGIVSLISLA ILSYDRYSTL 141
Am_SF_tmt6          GTSLSFAASV HGRWLVGRRG CMWYGFVNSC FGIVSLISLV VLSYERYSTL 150
Am_PC_tmt6          GTSLSFAASV HGRWLVGRRG CMWYGFVNSC FGIVSLISLV VLSYERYSTL 150
Am_SH_tmt6          GTSLSFAASV HGRWLVGRRG CMWYGFVNSC FGIVSLISLV VLSYERYSTL 150
Dr_tmt9             GTTLSFASSV RGRWLLGRHG CMWYGFINSC FGIVSLISLV VLSYDRYSTL 142
Am_PC_tmt9L         GTTLSFASSI HGRWLLGRQG CMWYGFINSC FGIVSLISLV ILSYDRYSTL 60
Am_PC_tmt9S         GTTLSFASSI HGRWLLGRQG CMWYGFINSC FGIVSLISLV ILSYDRYSTL 60
Dr_tmt14            GTPFSFASSL YGKWLLGHHG CKWYGFANSL FGIVSLMSLS ILSYERYAAL 146
Am_PC_tmt14         ---------- ---------- ---------- ---------- ---------- 0
Am_SH_tmt14         ---------- ---------- ---------- ---------- ---------- 0
Dr_tmt24            GTPFSFAAST QGRWLIGRAG CVWYGFINSF LGVVSLISLA VLSYERYCTM 140
Am_PC_tmt24         -----XAAST RGRWLIGGAG CVWYGFVNSF LGIVSLISLA VLSYERYCTM 45
Clustal Consensus
```

```
                                                TMIV
Bt_ROD_OPSIN        CKPMSNFRFG ENHAIMGVAF TWVMALACAA PPLVGWSRYI PEGMQCSCGI 189
Dr_tmt6             T-VYNKRAPD YSKPLLAVGG SWLYSLFWTV PPLLGWSSYG LEGAGTSCSV 190
Am_SF_tmt6          T-VYNKQAPD YRKPLLAVGG SWLYSLLWTV PPLLGWSSYG LEGAGTSCSV 199
Am_PC_tmt6          T-VYNKQAPD YRKPLLAVGG SWLYSLLWTV PPLLGWSSYG LEGAGTSCSV 199
Am_SH_tmt6          T-VYNKQAPD YRKPLLAVGG SWLYSLLWTV PPLLGWSSYG LEGAGTSCSV 199
Dr_tmt9             T-VYHKRAPD YRKPLLAVGG SWLYSLIWTV PPLLGWSSYG LEGAGTSCSV 191
Am_PC_tmt9L         T-VYNKKGPD YRKPLLAVGG SWLYSVVWTV PPLLGWSSYG LEGAGTSCSV 190
Am_PC_tmt9S         T-VYNKKGPD YRKPLLAVGG SWLYSVVWTV PPLLGWSSYG LEGAGTSCSV 190
Dr_tmt14            LRATKADVSD FRRAWLCVAG SWLYSLLWTL PPFLGWSNYG PEGPGTTCSV 196
Am_PC_tmt14         ---------- ---------- --------XL PPFLGWSSYG PEGPGTTCSV 22
Am_SH_tmt14         ---------- ---------- --------XL PPFLGWSSYG PEGPGTTCSV 22
Dr_tmt24            MGSTQADSTN YRKVVIGIAF SWIYSMVWTL PPLFGWSCYG PEGPGTTCSV 190
Am_PC_tmt24         MGATQADSTN YRKVIMGITF SWIYSMIWTL PPLFGWSRYG PEGPGTTCSV 95
Clustal Consensus                                   **:-*** *   **    :*-:

                                                TMV          Third cytoplasmic loop
Bt_ROD_OPSIN        DYYTPHEETN NESFVIYMFV VHFIIPLIVI FFCYGQLVFT VKEAAAQQQE 239
Dr_tmt6             TWTA--NTPQ SHSYIICLFI FCLGIPVLVM VYCYSRLICA VKQVGRIRK- 236
Am_SF_tmt6          TWTS--KTVQ SHSYIICLFI FCLGIPVLIM MYCYSRLLCA VKQVGRFRK- 245
Am_PC_tmt6          TWTS--KTLQ SHSYIICLFI FCLGIPVLIM MYCYSRLLCT VKQVGRFRK- 245
Am_SH_tmt6          TWTS--KTLQ SHSYIICLFI FCLGIPVLIM MYCYSRLLCA VKQVGRFRK- 245
Dr_tmt9             SWTQ--RTAE SHAYIICLFV FCLGLPVLVM VYCYGRLLYA VKQVGKIRK- 237
Am_PC_tmt9L         SWTE--HSPQ SHAYIICLFI FCLALPVLVM VYSYGRLLYA VKQVGKIRK- 155
Am_PC_tmt9S         SWTE--HSPQ SHAYIICLFI FCLALPVLVM VYSYGRLLYA VKQVGKIRK- 155
Dr_tmt14            QWHL--RSTS SISYVMCLFI FCLLLPLVLM IFCYGKILLL IKGVTKINL- 242
Am_PC_tmt14         QWHQ--RSPS SISYVVCLFI FCLVLPLLLM VYCYGKILFI IKGVTKINL- 68
Am_SH_tmt14         QWHQ--RSPS SVSYVVCLFI FCLVLPLLLM VYCYGKILFI IKGVTKINL- 68
Dr_tmt24            NWAA--RTPN NVSYIVCLFV FCLILPFIVI VYSYGRLLQA ITQVSRINT- 236
Am_PC_tmt24         NWAA--KTAN NVSYIICLFF FCLILPFFVI IYSYGKLLQA IKQVSRINT- 141
Clustal Consensus    :    -  - - :::: :*- - : :*--:: -:-*-:::   :- -   -

                                                TMVI                    TMVII
Bt_ROD_OPSIN        SATTQKAEKE VTRMVIIMVI AFLICWLPYA GVAFYIFTHQ GSDFGPIFMT 289
Dr_tmt6             -TAARREYH ILFMVITTVV CYLLCWMPYG VVAMMATFGR PGIISPIASV 286
Am_SF_tmt6          -TAARREYH ILFMVVTTVV CYLVCWMPYG VVAMTATFGQ PGLITPVVSV 295
Am_PC_tmt6          -TAARREYH ILFMVVTTVV CYLVCWMPYG VVAMTATFGR PGLITPVVSV 295
Am_SH_tmt6          -TAARREYH ILFMVVTTVV CYLVCWMPYG VVAMTATFGR PGLITPVVSV 295
Dr_tmt9             -TAARKREYH VLFMVITTVV CYLLCWMPYG VVAMMATFGR PGIISPVASV 287
Am_PC_tmt9L         -SAARREYH VLFMVITTVV CYLLCWTPYS VVAIMATFGR PGIITPVASV 205
Am_PC_tmt9S         -SAARREYH VLFMVITTVV CYLLCWTPYS VVAIMATFGR PGIITPVASV 205
Dr_tmt14            -LTAQRRENH ILLMVVTMVS CYLLCWMPYG VVALLATFGR TGLITPVTSI 292
Am_PC_tmt14         -LSAQRRENH ILLMVLTMVS CYLLCWMPYG VVSLMATFGK QGLITPVASV 118
Am_SH_tmt14         -LSAQRRENH ILLMVLTMVS CYLLCWMPYG VVSLMATFGK QGLITPVASV 118
Dr_tmt24            -VVSRKREQR VLFMVVTMVV CYLLCWLPYG IMALLATFGH PGLVTPAASI 186
Am_PC_tmt24         -VVSRKREQR VLLLVITMVV CYLLCWLPYG IMALVATFGH PGLVTPEASI 191
Clustal Consensus      ::: * - :  :*:  *  -:*:** **-  : :     :  - - *
```

```
Bt_ROD_OPSIN        IPAFFAKTSA VYNPVIYIMM NKQFRNCMVT TLCCGKNPLG DDEASTTVSK 339
Dr_tmt6             VPSLLAKSST VINPLIYILM NKQFYRCFLI LIHCKHSSLE NGQSSMPSRT 336
Am_SF_tmt6          VPSLLAKSST VFNPIIYILM NKQFYRCFLI LFHCKHSSHL NGH-SMPSRT 344
Am_PC_tmt6          VPSLLAKSST VFNPIIYILM NKQFYRCFLI LFHCKHSSHL NGH-SMPSRT 344
Am_SH_tmt6          VPSLLAKSST VFNPIIYILM NKQFYRCFLI LFHCKHSSHL NGH-SMPSRT 344
Dr_tmt9             VPSLLAKSST VINPLIYILM NKQFYRCFRI LFCCQRSLLQ NGHSSMPSKT 337
Am_PC_tmt9L         VPSLLAKSST VINPVIYILM NKQFYRCFLI LFHCRGRSME NGHSSMQSKT 255
Am_PC_tmt9S         VPSLLAKSST VINPVIYILM NKQVRAFIHS PKTNQSLIF- IYF*------ 240
Dr_tmt14            VPSVLAKSST VVNPVIYVLF NNQFYRCFVA FLKCQGEPSV HGQNPQHSSK 342
Am_PC_tmt14         VPSVLAKSST VINPVIYVLF NNQFYRCFIA FVKCGAEPSI QTLNPPHNSK 168
Am_SH_tmt14         VPSVLAKSST VINPVIYVLF NNQFYRCFIA FVKCGAEPSI QTLNPPHNSK 168
Dr_tmt24            VPSLLAKSST VINPIIYIFM NKQFCRCFHA LIMCTTPERG SSFKNSSKVT 336
Am_PC_tmt24         VPSLLAKSST VINPVIYIF- ---------- ---------- ---------- 210
Clustal Consensus   :*:-:**:*: * **:**::

Bt_ROD_OPSIN        TETSQVAPA* ---------- ---------- ---------- ---------- 349
Dr_tmt6             TGIQLNRRPY SNPVADNAPP SID------- ---LQN-DCS TPVSG*---- 371
Am_SF_tmt6          TVIQLNRRLC SNTVTGNIPT SLR------- ---LN-TEIS TPVSDRTHPP 379
Am_PC_tmt6          TVIQLNRRLC SNTVTGNIPT SLR------- ---LN-TEIS TPVSDRTHPP 379
Am_SH_tmt6          TVIQLNRRLC SNTVTGNIPT SLR------- ---LN-TEIS TPVSDRTHPP 379
Dr_tmt9             TVIQLNRRVN SNAVACTAQI STG------- ---THNHDCS THVTERSNPP 373
Am_PC_tmt9L         TVIHLNRVVH NNTVACQAQI STG------- ---LQEQGLS TTAMERSNPP 291
Am_PC_tmt9S         ---------- ---------- ---------- ---------- ----------
Dr_tmt14            EDPHVFRPCD GASIHRSAEG PQK------- ------KEQH SLSLVVHYTP 380
Am_PC_tmt14         EEAPASSFCR PWS--HSKL- SPA------- ------RECR TLSLVVHYTP 202
Am_SH_tmt14         EEAPASSFCR PWS--HSKL- SPA------- ------RECR TLSLVVHYTP 202
Dr_tmt24            KTLRTVRRAN GQNVTFAVAS AVHRTPYSDR QKSSSEGEKL PPATGQGTSK 386
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Bt_ROD_OPSIN        ---------- --
Dr_tmt6             ---------- --
Am_SF_tmt6          QITP*----- -- 388
Am_PC_tmt6          QITP*----- -- 388
Am_SH_tmt6          QITP*----- -- 388
Dr_tmt9             EVIP*----- -- 382
Am_PC_tmt9L         EISS*----- -- 300
Am_PC_tmt9S         ---------- --
Dr_tmt14            *--------- -- 380
Am_PC_tmt14         *--------- -- 203
Am_SH_tmt14         *--------- -- 203
Dr_tmt24            PVVSLVAYYN G* 398
Am_PC_tmt24         ---------- --
Clustal Consensus
```

(B)

		Rh1	**tmt6**		**tmt9**			**tmt14**		**tmt24**	
Conserved features	**AA position**	***Bt***	***Dr***	***Am***	***Dr***	***Am* long**	***Am* short**	***Dr***	***Am***	***Dr***	***Am***
Putative glycosylation sites	Varies	2	3	2	2	NA	NA	2	NA	2	NA
Conserved Glu or Tyr	113	E	Y	Y	Y	Y	Y	Y	Y	Y	NA
Conserved Cys pair	118, 187	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
E(D)/R/Y(W/F) motif	134-136	ERY	DRY	ERY	DRY	DRY	DRY	ERY	ERY	ERY	NA
Conserved Glu	181	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NPxxY(x)$_{5,6}$F motif	203-313	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Retinal binding site	296	K	K	K	K	K	K	K	K	K	K
Putative palmitoylation sites	322, 323	✓	✓	✓	✓	✓	✓	✓	✓	✓	NA
Potential phosphorylation sites * [threonine, serine]	Varies	9	12	17	15	10	3	8	11	19	NA

* Numbers of potential glycosylation and phosphorylation sites are indicated.

Figure 7.2 (A) Alignment of the amino acid sequences of *A. mexicanus* (*Am*) tmt opsins with zebrafish (*Danio rerio*, *Dm*) tmt opsins and bovine (*Bos taurus*, *Bt*) rod opsin. The different classes of tmt opsins are colour coded: class I (blue), class II (red), and class III (green). Gaps are inserted as dashes (-) to maintain a high level of identity between the sequences. Predicted transmembrane domains (TMs) are indicated by red and green lines above the residues. The putative positions of TMs for each tmt-opsin were predicted online using TMHMM Server Version 2.0 (http://www.cbs.dtu.dk/services/TMHMM/), and these were highlighted yellow. TMs shown for bovine rod opsin were determined from the published crystal structures (Palczewski et al., 2000). The first, second, and third cytoplasmic loops known to be involved in determining G protein specificity are highlighted with a box around the sequences. Symbols below each set of the alignment delineates level of conservation between the sequences: * = consensus residue between all 6 tmt opsin and bovine rod opsin sequences; := different aa but of similar charges across the sequences; .= different aa with distinctive charges across the sequences. Residues identified as essential for opsin structure are underlined and bolded, the predicted palmitoylation sites are highlighted light blue, and the predicted phosphorylation sites are shaded in grey, all of which are discussed in the text. **(B)** Table summarising the conserved structural and functional motifs of the *A. mexicanus* tmt opsins in comparison with the zebrafish tmt opsins and bovine Rh1.

Like many other non-visual opsins, such as zebrafish tmt opsins (Moutsaki et al., 2003), opn4 (Provencio et al., 2000), opn5 (Tarttelin et al., 2003), peropsin (Sun et al., 1997) and invertebrate visual pigments (Bellingham et al., 1998), the tmt opsins discovered from *A. mexicanus* thus far all contained an uncharged tyrosine residue at position 113. As predicted for these non-visual opsins, it is likely that their counterion is located at position E181 instead, which is also conserved in all the *A. mexicanus* tmt opsins examined in this study. The intracellular domains of GPCR are implicated for determining protein subtype specificity (refer to Section 6.4). A comparison of CL1, CL2, and CL3 of *A. mexicanus* tmt opsins with those of *D. rerio* tmt opsins shows a high degree of identity (87-100%) with their closely related zebrafish tmt opsin orthologues (see Table 7.2). Thus, it may be deduced that *A. mexicanus* tmt opsins are likely to couple to the same G-protein signalling pathway (G_i-type) as zebrafish tmt opsins. Evidence supporting this comes from the analysis of full-length *A. mexicanus* tmt6 sequences using Pred-Couple v1.0 web server (http://athina.biol.uoa.gr/bioinformatics/PRED-COUPLE/), which predicted these opsins to couple with G_i-type G proteins.

Table 7.2 *Sequence alignments of the intracellular loops of A. mexicanus (Am- SF, PC, SH) tmt opsins with D. rerio (Dr) tmt opsins.*

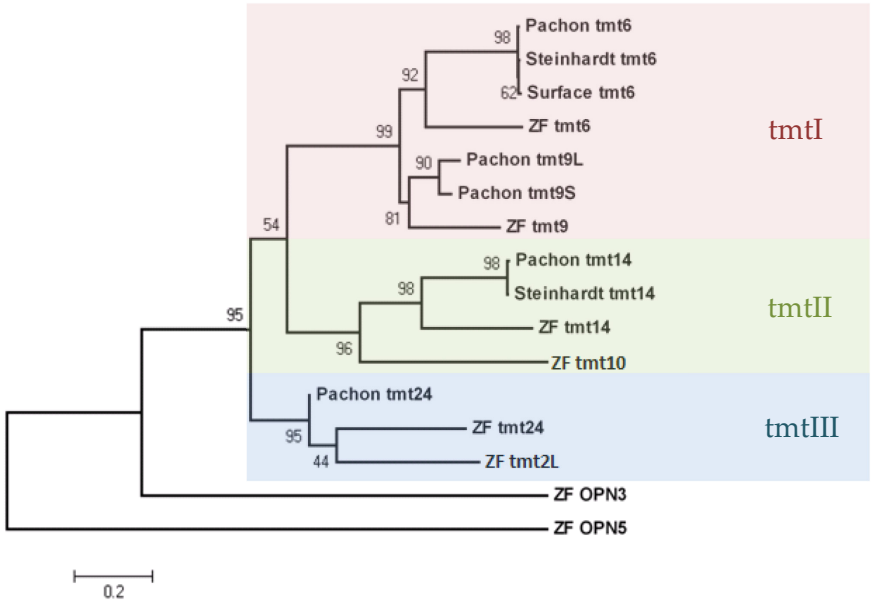
Cytoplasmic domains	G-protein	Opsins	Sequence alignment	% similarity to their respective orthologues in zebrafish
Loop 1		Dr tmt6	FCKFKTLRTPVN	100
		Am (SF) tmt6	FCKFKTLRTPVN	100
		Am (PC) tmt6	FCKFKTLRTPVN	100
		Am (SH) tmt6	FCKFKTLRTPVN	100
Loop 2		Dr tmt6	DRYSTLTVYNKR~APDYSK	100
		Am (SF) tmt6	ERYSTLT~VYNKQAPDYRK	91
		Am (PC) tmt6	ERYSTLT~VYNKQAPDYRK	91
		Am (SH) tmt6	ERYSTLT~VYNKQAPDYRK	91
		Dr tmt9	DRYSTLTVYHKR~APDYRK	100
		Am (PC) tmt9l	DRYSTLT~VYNKKGPDYRK	91
		Am (PC) tmt9s	DRYSTLT~VYNKKGPDYRK	91
		Dr tmt24	ERYCTMMGSTQADSTNYRK	100
		Am (PC) tmt24	ERYCTMMGATQADSTNYRK	95
Loop 3		Dr tmt6	SRLICAVKQVGRIRKTAARRREY	100
		Am (SF) tmt6	SRLLCAVKQVGRFRKTAARRREY	91
		Am (PC) tmt6	SRLLCAVKQVGRFRKTAARRREY	91
		Am (SH) tmt6	SRLLCAVKQVGRFRKTAARRREY	91
		Dr tmt9	GRLLYAVKQVGKIRKTAARKREY	100
		Am (PC) tmt9l	GRLLYAVKQVGKIRKSAARRREY	91
		Am (PC) tmt9s	GRLLYAVKQVGKIRKSAARRREY	91
		Dr tmt14	GKILLLIKGVTKINLLTAQRREN	100
		Am (PC) tmt14	GKILFIIKGVTKINLLSAQRREN	87
		Am (SH) tmt14	GKILFIIKGVTKINLLSAQRREN	87
		Dr tmt24	GRLLQAITQVSRINTVVSRKREQ	100
		Am (PC) tmt24	GKLLQAIKQVSRINTVVSRKREQ	91
C-terminus		Dr tmt6	ILMNKQFYRCFLILIHCKHSSLENGQSSMPSRTTGIQLNRRPYSNPVADNAPPSIDLQNDCSTPVSG*	100
		Am (SF) tmt6	ILMNKQFYRCFLILFHCKHSSHLNGHS~MPSRTTVIQLNRRLCSNTVTGNIPTSLRLNTEISTPVSDRTHPPQITP*	61
		Am (PC) tmt6	ILMNKQFYRCFLILFHCKHSSHLNGHS~MPSRTTVIQLNRRLCSNTVTGNIPTSLRLNTEISTPVSDRTHPPQITP*	61
		Am (SH) tmt6	ILMNKQFYRCFLILFHCKHSSHLNGHS~MPSRTTVIQLNRRLCSNTVTGNIPTSLRLNTEISTPVSDRTHPPQITP*	61
		Dr tmt9	ILMNKQFYRCFRILFCCQRSLLQNGHSSMPSKTTVIQLNRRVNSNAVACTAQISTGTHNHDCSTHVTERSNPPEVIP*	100
		Am (PC) tmt9l	ILMNKQFYRCFLILFHCRGRSMENGHSSMQSKTTVIHLNRVVHNNTVACQAQISTGLQEQGLSTTAMERSNPPEISS*	65
		Am (PC) tmt9s	ILMNKQVR~~~~AFIHSP~~~~~~~~~~~KTNQSLIFIYF*	0
		Dr tmt14	VLFNNQFYRCFVAFLKCQGEPSVHGQNPQHSSKEDPHVFRPCDGASIHRSAEGPQKKEQHSLSLVVHYTP*	100
		Am (PC) tmt14	VLFNNQFYRCFIAFVKCGAEPSIQTLNPPHNSKEEAPASSFCR~~PWSHSKLSPAR~ECRTLSLVVHYTP*	54
		Am (SH) tmt14	VLFNNQFYRCFIAFVKCGAEPSIQTLNPPHNSKEEAPASSFCR~~PWSHSKLSPAR~ECRTLSLVVHYTP*	54

Table 7.2 Sequence comparison of the intracellular domains of *A. mexicanus* tmt opsins with those of *D. rerio*. The opsin sequences were aligned using ClustalW (Higgins et al., 1996).

7.2.3. Phylogenetic analyses of *A. mexicanus* tmt opsins

Phylogenetic analyses show that the *A. mexicanus* tmt opsins are part of the multiple tissue opsin group that includes opn3 and opn5 (Fig. 7.3). As seen from the phylogenetic trees, the nucleotide and amino acid sequences of *A. mexicanus* tmt opsins form a single clade with those of the zebrafish tmt opsins, which is separated from their most closely related sister group opn3. The topology of the nucleotide and amino acid trees resemble each other, and the phylogenetic positioning of most tmt opsins are supported by high bootstrap values (98% - 100%). The class-specific clustering of zebrafish and *A. mexicanus* tmt opsins is also well supported in the bootstrap analysis, though the precise position of surface tmt6 and zebrafish tmt24 cannot be established due to their poor bootstrap support, only 61% and 38% respectively. The presence of tmt opsin orthologues from each of the tmt classes in *A. mexicanus* provides further evidence supporting the occurrence of whole genome duplication in the common ancestor of teleosts (Taylor et al., 2001, Meyer and Schartl, 1999, Hoegg et al., 2004).

(A)



(B)

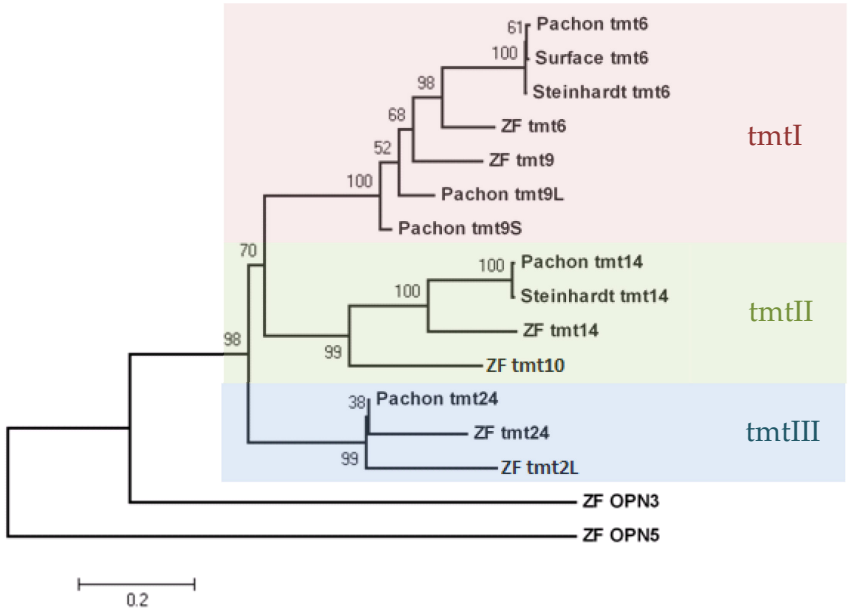


Figure 7.3 Phylogenetic analyses of zebrafish and *A. mexicanus tmt* opsins. The trees were constructed on codon-matched sequence alignments of **(i)** the nucleotide sequences using Maximum Composite Likelihood (MCL) (Tamura and Nei, 1993), and **(ii)** the amino acid sequences using Jones-Taylor-Thornton (JTT) substitution matrix (Jones et al., 1992). Neighbour-joining trees were generated with 1000 bootstrapping replicates. Internal branch supports are expressed as percentages at the node of each branch. The scale bars at the bottom of figures denote the number of nucleotide or amino acid substitutions per site.

7.2.4. Spectral property of tmt6 from cave-form (Pachón) Mexican tetra

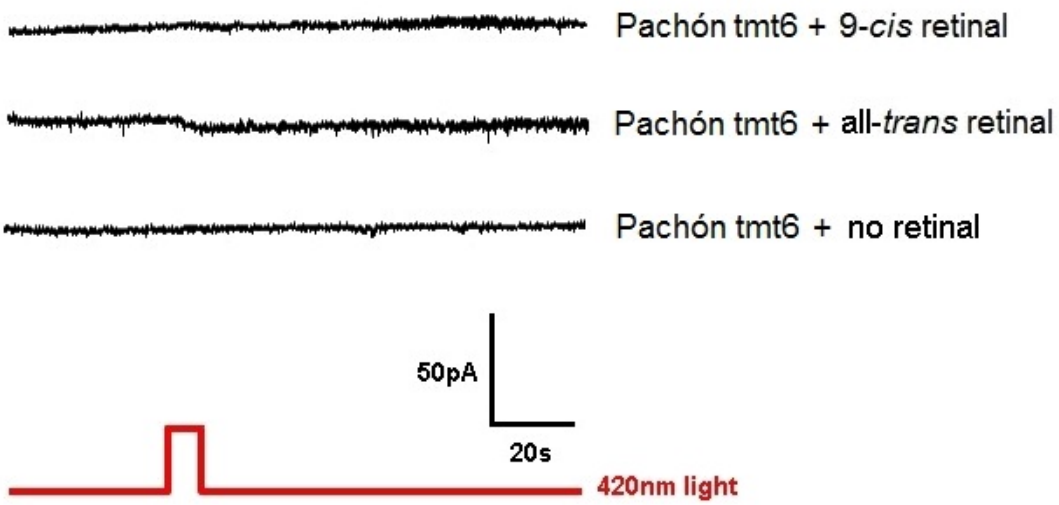
A representative tmt opsin from cave-form *A. mexicanus*, Pachón tmt6, was characterized spectrally by *in vitro* spectrophotometric analysis. This tmt opsin was chosen for investigation over the one from epigean form, as it was of interest to see if the tmt opsin retained light-sensing ability over the course of many generations in cave environment. Since sequence analyses showed that tmt6 from the hypogean *A. mexicanus* populations share a high identity (99.2%) with that of the epigean population, they are likely to have common tertiary structure and possible light-sensing function. Using the same protocol as described in Sections 5.4.1 and 5.4.4, dark and acid-denatured spectra for each reconstituted pigment of tmt6 with 11-*cis* and all-*trans* retinal were measured over a broad range of wavelengths from 200 nm to 700 nm.

Unfortunately, with both 11-*cis* and all-*trans* retinal, Pachón tmt6 failed to regenerate as a photoactive pigment *in vitro* (data not shown). This result is similar to the tmt6 of *D. rerio*, which suggest that it is likely to be a technical issue. It would be useful to test the photosensitivity of other *A. mexicanus* tmt opsins, e.g. tmt10, though their full-length sequences have not yet been cloned due to time limitation.

7.2.5. G protein activation by Pachón tmt6

The G_q-type G protein signalling cascade of Pachón tmt6 was investigated by whole-cell patch clamp recording as a comparative, functional analysis with zebrafish tmt6 (see Section 6.3.3 for detailed methodology). Figure 7.4A showed the current traces recorded from the opsin-transfected and untransfected cells in the absence or presence of retinal (9-*cis* and all-*trans*). The results demonstrated that although retinal-dependent light responses were detected in tmt-opsin transfected cells, they were not significantly greater than the background response measured from the untransfected control (Fig. 7.4B). Hence, these were likely intrinsic responses mediated by endogenous opsins and not tmt opsins in the Neuro-2A cells.

(A)



(B)

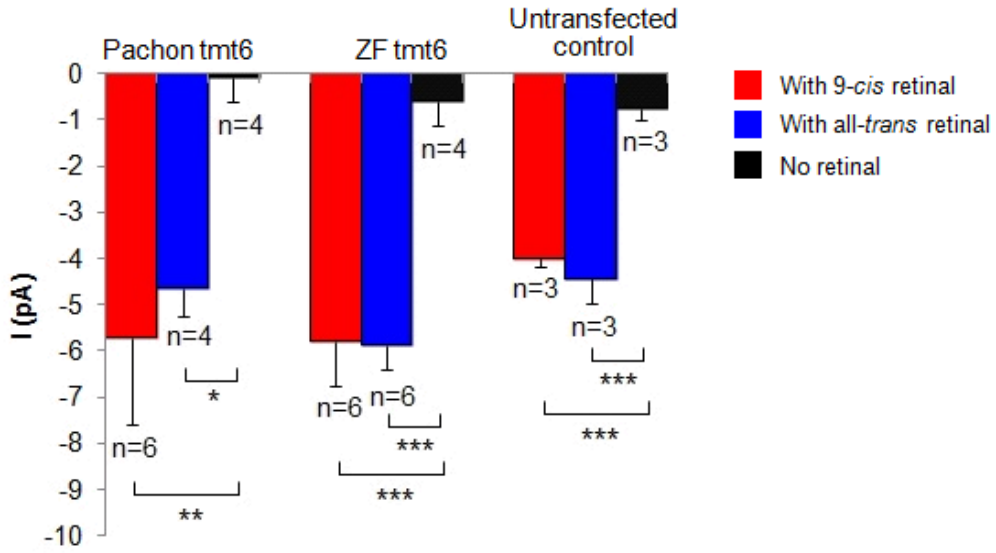


Figure 7.4 Light-evoked responses in Neuro-2A cells expressing Pachón tmt6. **(A)** Whole-cell currents recorded from cells expressing Pachón tmt6, pre-incubated with 9-*cis* retinal, all-*trans* retinal, and without any retinal. **(B)** Peak current amplitudes measured from the opsin-transfected cells upon light pulse stimulation with 9-*cis* retinal (red bars), all-*trans* retinal (blue bars), and no retinal (black bars). Baselines for background light-evoked responses due to the presence of intrinsic opsins are set at -4 ± 0.17 pA for 9-*cis* retinal and -4.43 ± 0.54 pA for all-*trans* retinal based on recordings from untransfected cells. Data were analysed using student's *t* test for the parametric data. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

7.3 Discussion

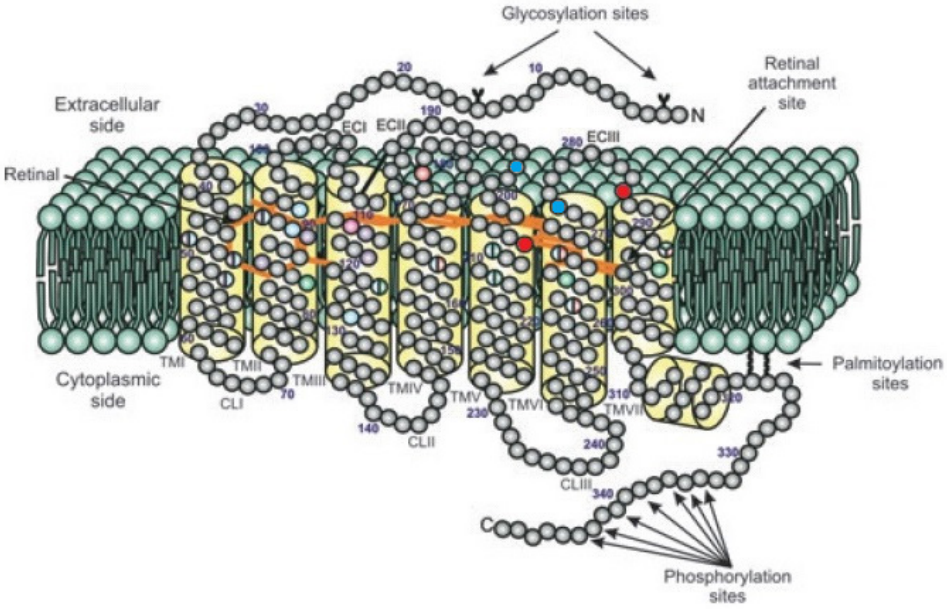
This present study describes the identification of tmt opsins in surface and cave-dwelling populations (Pachón and Steinhardt) of *A. mexicanus*, which are orthologous to the tmt opsins discovered in zebrafish (Chapter 3). RT-PCR analysis demonstrated that in the *A. mexicanus* cell lines derived from embryos of these three different populations, at least one of the tmt opsins from each class- *tmt6* and *tmt9* (tmtI), *tmt14* (tmtII), and *tmt24* (tmtIII) exists (Section 7.2.1). *tmt6* is thus far the only full-length sequence obtained, whilst partial sequences of the other tmt transcripts were confirmed. Although partial sequences may generate functional or non-functional proteins in the cells, their expression in the embryonic tissues together sequence motif analyses showed that the opsins are likely to be intact, suggesting that they are functional. It is possible that expression of other *tmt* opsin members (e.g. *tmt2* and *tmt10*) could be present in other *A. mexicanus* tissues but were not detectable in the embryonic cell lines. Future experiments can be done by PCR-based screenings of genomic DNA from the native, adult tissues of *A. mexicanus* for identification of other *tmt* opsin genes.

Since *A. mexicanus* tmt opsins possess K296, this suggests that they are able to bind retinal chromophore. The presence of a lysine in the seventh transmembrane segment at a location highly conserved across all known opsins indicates that *A. mexicanus* tmt opsins may covalently attach chromophore via a Schiff base linkage to the amino group of lysine (Terakita, 2005). Nonetheless, it is not known whether *A. mexicanus* tmt opsins function as a light-sensing photopigment, a photoisomerase, a retinal-binding protein, or has other atypical functions.

In preliminary studies, UV-visible absorption spectra of Pachón tmt6 showed that no absorption band was detected between 200 to 700 nm wavelengths, not even at the UV region (~370 nm) where maximum absorbance peaks of the zebrafish tmt opsins were identified (Section 5.2.1). The tmt opsins of *A. mexicanus* may bind other isomers of retinal chromophore (e.g. 7-*cis* or 13-*cis* retinal) which has not been tested. It is also possible that the retinal chromophore may only form complexes with tmt opsins during a specific stage of their generation in the native cells, which has been implicated in *Drosophila* visual pigments, as the binding may require cofactors (e.g. retinol dehydrogenase or arrestin) that are not present in the heterologous expression system (Ozaki et al., 1993, Kiselev and Subramaniam, 1997). Even though these reconstituted tmt6 opsins were predicted to undergo photochemical change upon light stimulation, they could equally likely form UV-absorbing, photostable pigments.

Despite being eyeless and living in perpetual darkness, sequence analysis showed that both Pachón and Steinhardt cavefish have retained at least an intact *tmt6* in their genomes. A comparison between the epigean and the hypogean forms of tmt6 revealed two amino acid substitutions, V206L and Q285R, which are distant from the bound chromophore (Fig. 7.5). *A. mexicanus* tmt6 shares a high degree of identity in the amino acid sequence to zebrafish *tmt6* (~75%). Their corresponding locations in the zebrafish orthologues are also away from the attached retinal and are not in the intracellular regions, suggesting *A. mexicanus* tmt6 may function like the tmt6 in zebrafish.

(A)



(B)

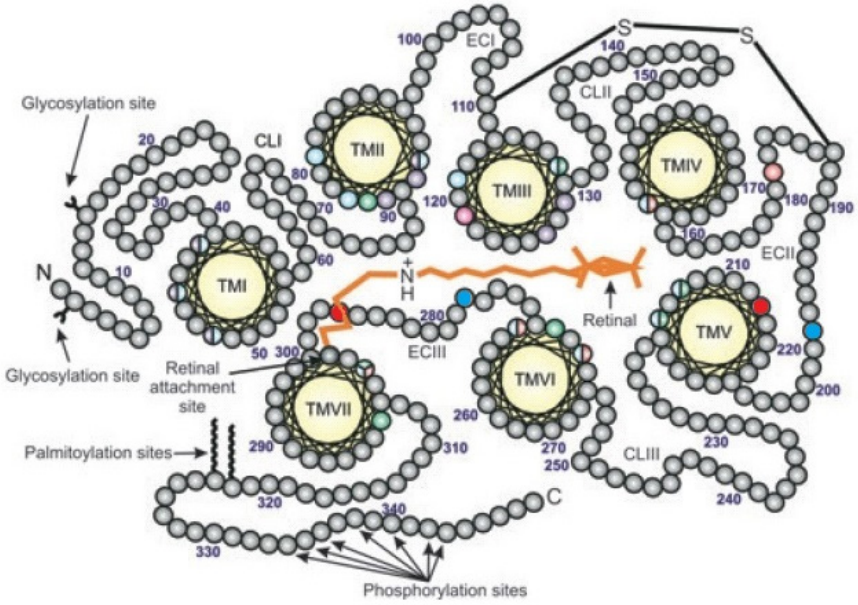


Figure 7.5 Representation of a typical opsin photopigment. **(A)** Side view and **(B)** membrane cross-section view of the model pigment showing seven transmembrane segments, extracellular loops (ECI, ECII, and ECIII), cytoplasmic loops (CLI, CLII, and CLIII), amino (N)- and carboxyl (C)- termini, and a bound retinal chromophore (diagrams modified from Davies et al., 2012). Distinctive amino acid residues between tmt6 of the hypogean *A. mexicanus* (Pachón and Steinhardt) and the epigean form are highlighted in red. The locations of these resides in their corresponding *D. rerio* orthologue are highlighted in blue.

Given the negative data from electrophysiological studies for zebrafish tmt opsins, failure of the Pachón tmt6 indicates that it was probably due to technical issues. Further experiment would need to be performed to identify its potential G-protein signalling pathway using different assays, e.g. bioluminescent reporter assays, though this has not been done due to time constraints. Nonetheless, this raises intriguing questions as to why these tmt opsins are maintained in the genome of the troglomorphic form of *A. mexicanus*, which diverged from the surface variant 1 to 3.1 million years ago (MYA) (Borowsky, 2008, Gross, 2012, Jeffery, 2009). Perhaps conserved linkage means that the gene is maintained, but its promoter might be altered and is no longer functional. Comparative analyses of the promoters of epigean and hypogean *A. mexicanus tmt6* can be performed to investigate the integrity of these regulatory regions, which may be linked to evolution of the cave form. Since cone opsin genes have been identified in the cave-dwelling form of *A. mexicanus*, this may suggest that it has not been long enough to lose the gene (Parry et al., 2003, Yokoyama and Yokoyama, 1990). There is a possibility that the mutation rate in the cavefish is slower than in the surface fish, which has been implicated from the lower genetic diversity measured by microsatellite typing in the cave populations compared to the surface populations (Strecker et al., 2003, Bradic et al., 2012).

Speculation on the potential selective motive for the maintenance of *tmt* opsins in *A. mexicanus* is difficult with limited spectrophotometric and functional data of the reconstituted pigments. Interestingly, orthologous of zebrafish *tmt2* and *opn4m-2* genes were identified in another species of cavefish from Somalia, *Phreatichthys andruzzii* (*P. andruzzii*) (Cavallari et al., 2011). Both the *P. andruzzii tmt2* and *opn4m-2*

genes were found to be mutated, with an insertion of a T at position 654 in *tmt* opsin and a deletion of a G at position 842 in *opn4* (the numbering is relative to the start codon of orthologous sequences in zebrafish) (Cavallari et al., 2011). The resulting truncated forms of the opsins lack the retinal binding region essential for their function as a light-sensing molecule. It was argued that these opsin mutations render *P. andruzzii* unable to detect light and thus photoentrain their peripheral circadian rhythms. Consequently, the circadian rhythmicity in *P. andruzzii* is regulated by other physiological cues, such as daily feeding time rather than the day-night cycle (Mejia, 2011, Cavallari et al., 2011). However, the study by Cavallari et al. (2011) does not explore the possibility that other *tmt* opsins may be present and the mutations might only occur in one *tmt* opsin gene.

In *A. mexicanus*, the techniques used also failed to detect other *tmt* opsins (e.g. *tmt10*). Nonetheless, both their hypogean and epigean forms express a *tmt6* that does not have insertion mutation or premature stop codon. Despite the explanations proposed above, it is not certain why these opsins did not regenerate *in vivo* with UV-vis spectrophotometric analysis. It may be that tmt6 does not act as a light-sensor but have different functional role, such as retinal-binding protein. Equally, it may reflect technical problems in generating functional proteins *in vitro*. More research will need to be conducted using high performance liquid chromatography (HPLC) analysis to confirm its interaction with retinal chromophores. Other *A. mexicanus* tmt opsins will also need to be tested to elucidate their potential light-sensitivity and their transduction cascades.

CHAPTER 8

General Discussion and Future Directions

CHAPTER 8: General Discussion and Future Directions

Majority of the zebrafish tissues are not only light sensitive (Tamai et al., 2005), but they also exhibit a highly decentralised circadian clock system that is directly photoentrainable (Whitmore et al., 1998, Whitmore et al., 2000, Kaneko et al., 2006). Although the identities of the photopigment(s) responsible for mediating these light responses observed are not known, tmt opsin has been suggested as a potential candidate (Moutsaki et al., 2003). The work described in this thesis has focused on identifying and characterising all the members of the *tmt* opsin multigene family in the zebrafish genome. Six zebrafish *tmt* opsin genes (*tmt2l (long)*, *tmt2s (short)*, *tmt6*, *tmt9*, *tmt10*, *tmt14* and *tmt24*) were isolated. Phylogenetic analyses showed that they clade into three gene subclasses (tmtI, tmtII, and tmtIII), forming a sister group to opn3 with ~87% amino acid identity. The RNA expression pattern of the *tmt* opsins in D1-D5 embryos and a panel of native adult tissues (brain, eye, testis, gut, gill, and fin) were examined by RT-PCR and NanoString analysis. With the cloning of representative zebrafish *tmt* opsins, their biochemical characteristics as a group of UV sensitive (UVS) photopigments has been elucidated. Using HEK293T and Neuro2A cells as heterologous expression systems to test the molecular signatures for each tmt opsin, initial insights were provided with regards to their spectral characteristics and their potential signal transduction pathway. However, it is important to recognise that these expression systems have limitations. Neither of the mammalian cell lines can recapitulate the exact physiological condition of zebrafish cells in which the tmt opsins were expressed. Although both cell lines have been

well-characterized and widely used for studying opioid receptor G-protein signalling (Spencer et al., 1997), the opsins expressed in the *in vitro* system may promiscuously bind to different G proteins or they might require specific cofactors during their biosynthesis for proper functions, as have been suggested for *Drosophila* rhodopsin (ninaE) which couples with chromophore 3-hydroxyretinal for maturation (Ozaki et al., 1993). Therefore, it is possible that the signalling pathway identified may not be the same one as in the native cells. Unfortunately, all zebrafish cell lines that have been studied are known to be intrinsically photosensitive (Whitmore et al., 2000) and some of them (e.g. PAC2 cell line) have already been shown to express a range of opsins including *tmt2* and *opn4* (Moutsaki et al., 2003, Farhat et al., 2009). Since multiple opsins are present in these zebrafish cell lines, it is difficult to elucidate the specific function of tmt opsins compared to the other opsin types. Thus, HEK293T and Neuro-2A cells were the most suitable cell lines chosen for this present project. These cell lines have the advantage of being immortal, providing a constant supply of the cells that can be genetically manipulated by *in vitro* transfection for functional studies. Neuro-2A cells have the ability to differentiate into neurons within three days, thus its signalling pathways closely mimic those of the matured neurons after incubation with retinoic acid, providing a useful tool for measurement of photocurrents. Both the Neuro-2A and HEK293T cells are known for appropriate membrane trafficking of opsin proteins (Jin et al., 2009, Chow et al., 2011, McClements et al., 2013), which allows for light stimulation testing of the transmembrane pigments. These cell lines therefore provided tools in which the

following questions were addressed regarding the structural and functional properties of zebrafish tmt opsins:

(1) Are zebrafish tmt pigments photosensitive? And if so, what are their spectral absorption peaks?

(2) Which physiologically-relevant chromophores are used by tmt opsins for light detection?

(3) Does light activation of zebrafish tmt pigments induce a cellular response in an *in vitro* mouse Neuro-2A cell model system? And if so, which G protein-coupled signalling pathway is activated?

(4) What are the mechanisms that underlie the spectral tuning of UV sensitivity in zebrafish tmt pigments?

8.1 Developmental and regional-specific expression of tmt opsins

In Chapter 4, the expression patterns of all six *tmt* opsin genes identified in the zebrafish genome were studied in both embryonic and adult zebrafish tissues using NanoString nCounter analysis and *in situ* hybridisation. Analysis of zebrafish embryos showed that tmt opsin transcripts were expressed early in development, with *tmt9* and *tmt24* being the first to be detected in day 1 embryos. From day 2 to day 5 after fertilization, *tmt* opsin expression was selectively localised in regions that are crucial for CNS development (e.g. in notochord and neural tube), formation of the digestive system (e.g. gut/liver), and control of autonomic functions (e.g. medulla oblongata in the hindbrain) (see Fig. 4.4). In addition, *tmt2* transcripts were detected in the neurocranial trabecula domain, which was not observed with other tmt opsins.

This trabecula has been implicated for structural function, forming a boarder to define morphogenetic regions within the brain (Cerny et al., 2006). The presence of *tmt* opsins could be important for transmitting light input signals to entrain the circadian rhythms in these regions. They might also be linked to other photosensitive cell pathways. It is possible that these tmt pigments may also act as UV-absorbing molecules to protect the areas from harmful UV irradiation. Taken together, findings from whole-mount *in situ* hybridisation demonstrated that the expression of all six *tmt* opsin genes are regulated both temporally and spatially in zebrafish embryos.

RNA *in situ* hybridisation studies of adult zebrafish ocular and brain tissues also revealed a distinctive expression pattern for each *tmt* opsin. Although all *tmt* opsins were prominently expressed in retinal bipolar cells, different *tmt* opsins may be localised to certain subtypes of the bipolar cells. To confirm the cell-type specific expression of *tmt* opsins, immunocytochemistry will need to be performed using flurophore-conjugated antibodies targeting different *tmt* opsins and differentiation of expression in certain cell types can be visualised under fluorescent microscopy. Some *tmt* opsins were detected in other retinal cell types, including the retinal pigment epithelium (*tmt9* and *tmt10*), amacrine cells (*tmt6*), and retinal ganglion cells (*tmt2*, *tmt9*, *tmt10* and *tmt14*). In the brain, the expression of *tmt* opsins in the olfactory bulb, preoptic tectum, thalamus, cerebellum, and the medullary gustatory column indicates that these genes may play diverse roles in feeding behaviour (Tabor et al., 2004, Hansen and Zeiske, 1998), motor sensory functions (Wagner, 2001), and vision (Nikolaou and Meyer, 2012). As the bistable nature of tmt pigments indicated that they may function as photopigments as well as photoisomerases, some of the tmt

pigments may function as G protein signal transducer, whilst others may act as photoisomerases to regenerate specific retinal isomers in the same cells. Although R-ISH was appropriate for identifying *tmt* opsin expression, there were some disadvantages. For detection of mRNA transcripts in the embryos, this method was limited by three-dimensional visualisation of the gene expression within the tissues. The experiment could have been performed on sections of embryos, but it was more important to see the overall gene expression in intact embryos, due to their small sizes and the low expression of *tmt* opsins during early development (as confirmed by NanoString data). R-ISH on adult tissue sections, like the brain and the retina, would allow for identification of the transcripts at cellular level and not just the gross anatomy. However, this technique would only allow for detection of a single gene at one time. In order to identify co-expression of *tmt* opsins in the cells, fluorescence *in situ* hybridisation (FISH) may be used, as two or more different probes labelled with distinct flurophore can be visualised at the same time (Bridger and Volpi, 2010). Additionally, *in situ* hybridisation only provides information on gene transcript expression and not protein expression, which can be influenced by post-translational modification. In order to confirm the presence of tmt opsin protein in these tissues, one can perform immunohistochemistry (IHC) using antibodies specifically designed to bind to the proteins via a unique epitope.

5′-RACE PCR was performed as a prerequisite to obtain the 5′-end of *tmt* opsin sequence for promoter analysis. Regulation of complex *tmt* opsins gene expressions was investigated by examining these promoter regions. From promoter analysis of *tmt2* and *tmt6* opsins, the study revealed a number of *cis*-acting elements

that correlates with the expression profile observed. For example, all *tmt* opsins are expressed in bipolar cells, and both *tmt2* and *tmt6* possess multiple chx10 elements that have been implicated in the development of these cells (Dorval et al., 2006). The presence of pax4 and pax6 elements indicated that these genes may have a role in ocular development, which agrees with their expression in the retinal tissues. D-box and e-box elements (Munoz and Baler, 2003, Zhang et al., 2004) were also identified in the promoters of *tmt2* and *tmt6*, suggesting their potential roles in controlling central or peripheral circadian clocks. A homeobox gene nkx2.5 essential for heart development (Balci and Akdemir, 2011, Guner-Ataman et al., 2013) was present in *tmt6* promoter, which indicates that it may have a role linked to cardiac functions. Collectively, these data are in agreement with the broad expression pattern of *tmt* opsins in zebrafish. Identification of the clock-related regulatory elements at the promoters of these *tmt* opsins indicates that their expression might be controlled in a circadian cycle.

8.2 Zebrafish tmt opsins function as light sensing molecules

Several lines of evidence from this project have strongly demonstrated that the zebrafish tmt opsins can form photosensitive molecules. The deduced amino acid sequence of each zebrafish tmt opsin in Chapter 3 adheres to the secondary structure of other known GPCRs, containing all of the conserved residues critical for correct protein folding and functional expression. These opsin sequences also possess Lys296, which is a marker of opsin pigments (Hargrave et al., 1983). As shown in Chapter 5, tmt10 was expressed successfully *in vitro* and reconstituted with both 11-

cis and all-*trans* retinals to give absorption maxima in the UV region between 365 – 370 nm.

For visual pigments, it has been shown that the covalent Schiff base linkage between the retinal chromophore and the opsin protein contributes to the spectral tuning of photopigments (Kusnetzow et al., 2004). Amongst the visual opsins, UVS cone pigments, encoded by the *SWS1* gene, are the only photopigments with an unprotonated Schiff base in their dark state (Kono, 2006). Kusnetzow and colleagues postulated that a counterion switching occurs from Glu113 to Glu181 during photoactivation that leads to a second protonation event of the Schiff base when these cone pigments are irradiated by UV. Based on this hypothesis, pH dependency of a representative tmt pigment (tmt10) was investigated. The results in Chapter 5 showed that the λ_{max} of the reconstituted tmt10 with 11-*cis* retinal remained at ~370 nm whether the pH of its surrounding was increased or decreased, suggesting that the pigment is already protonated in the dark. It could also be argued that the tmt10 pigment might be unprotonated, as increasing the pH to 9.5 did not cause a spectral shift towards even shorter wavelength. However, results from the mutation studies provided further evidence that tmt10 pigment is more likely to be protonated, as addition of a counterion at site 113 did not lead to a spectral shift towards the violet region (see Section 5.2.3). Previous studies have shown that amino acid changes at site 86 also play a role in the spectral tuning of SWS1 pigments between violet (385-445 nm) and UV (355-365 nm) sensitivities (Cowing et al., 2002, Hunt et al., 2004, Carvalho et al., 2007). Sites 90 and 93 are also important for spectral tuning in birds, as all their UVS pigments are found to have a Cys90, whereas the VS pigments

possess Ser90 (Yokoyama et al., 2000). All zebrafish tmt opsins have a cysteine residue at site 87, which is located close to the above mentioned sites that have been implicated for spectral tuning in UVS pigments. It was hypothesised that the presence of cysteine residue at position 86 in non-avian SWS1 pigment imposes steric hindrance in the opsin molecule which prevents its Schiff base from being protonated in the dark (Shi and Yokoyama, 2003). However, when a Cys87Ser substitution was introduced in zebrafish tmt10, its λ_{max} remained in the UV region at 369 nm, suggesting that the Cys87 has no spectral tuning effect. Findings from both pH-dependency experiments and mutational studies indicate that zebrafish tmt10 is likely to be a protonated photopigment in the dark. It is likely that they employ a novel UV spectral mechanism which has yet to be identified. Further mutational studies testing different amino acid sites surrounding the retinal binding pocket, such as position 90 and 93 (Hunt et al., 2004, Odeen and Hastad, 2009), which have also been implicated for spectral tuning in avian UVS pigments, will be required to elucidate the spectral tuning of zebrafish tmt opsins. As these sites are important for avian SWS1 tuning, the effect on zebrafish photopigments (e.g. tmt pigments) is unclear. As mentioned before, a pufferfish tmt pigment and mouse Opn3 with 9- and 11-*cis* retinals have been shown to have their respective λ_{max} measured at ~460 nm and ~500 nm in the dark. These spectral peaks of *T. rubripes* tmtII and Opn3 determined at 4°C are significantly longwave shifted compared to those of the zebrafish tmt pigments measured at room temperature (Koyanagi et al., 2013). Although temperature variation could to be a contributor to the spectral differences recorded, different amino acid at specific spectral tuning sites may also have a

significant influence. It would be interesting to compare the sequence of pufferfish *tmt* opsin with those of zebrafish *tmt* opsins in detail for identification of putative spectral tuning sites that can be tested by mutagenesis.

Functional expression studies in Chapter 6 provided further evidence that zebrafish tmt opsins are capable of mediating light-dependent cellular responses in the heterologous HEK293 cell expression system. Using a luciferase-based bioluminescence light response assay (Bailes et al., 2012), it was revealed that tmt6, tmt10, tmt14 and tmt24 specifically couple to a G_i-type (but not G_s and G_q-type) G protein signalling cascade, which led to an inhibition of cAMP responses in transfected cells. Results from Ca^{2+} imaging and whole-cell patch clamp recordings using Neuro2A cells also indicated that none of the tmt opsins binds significantly to G_q- or G_s-type pathways. However, as mentioned previously, the heterologous cell systems employed in this project might not contain the essential machinery of native zebrafish cells that are required for the *in vivo* function of tmt pigments. Therefore, a possible experimental approach that may confirm the G protein interaction of tmt opsins is to use a G protein activation assay (Terakita et al., 2002). This assay detects guanosine diphosphate (GDP)/guanosine triphosphate (GTP)γS exchange reaction by G-protein based on quantification of radioactive [^{35}S] GTPγS using a liquid scintillation counter. Unfortunately, our laboratory was not set up for this assay.

In light of the present data pertaining to tmt opsin-driven activation of a G_i-signalling cascade, the resulting suppression of intracellular cAMP response may be linked to numerous biological functions ranging from circadian clock regulation, survival of neurons, to biosynthesis of neurotransmitters (as discussed in Chapter 6).

To examine the physiological role(s) of each tmt opsin, targeted *tmt* opsin gene expression can be knocked down using antisense morpholino oligonucleotides. Morpholino oligonucleotides are short oligos (usually 25 nucleotides) designed to bind to specific sequences of the tmt opsin mRNA by complementary base pairing, and thus blocking the access of protein complexes required for translation (Heasman, 2002). Administration of these morpholino oligos can be achieved by microinjection into zebrafish embryos, and phenotypic effects of the targeted gene inhibition can be studied during their early developmental period. Physiologies of the animals that can be measured from the gene silencing effect may include level of unbound retinal in the cells, their circadian physiologies (e.g. in sleep, metabolism and their motor functions), survival of neurons in response to UV irradiation, production of neurotransmitters, or specific developmental processes such as maturation of bipolar cells. Although morpholino-based screening has shown to be achievable in zebrafish, the experimental design has to be thoughtfully handled, tested for toxicity, and validated for gene-specificity prior to use. With the presence of multiple *tmt* opsin genes, one potential problem is that there may be some confounding redundancy such that when the expression of one gene is removed, others can replace its function. Other techniques employing artificially engineered nucleases e.g. zinc finger nucleases (ZFNs) and TAL effector nucleases (TALENs) can also be used to control gene expression in an organism (Kim et al., 2013). These nucleases possess two domains, a customizable DNA-binding domain for recognition of the targeted sequence and an unspecific nuclease domain that cleave the DNA. Genetic modification mediated by these nucleases is advantageous in that they are highly

specific for their binding sequences and can be designed to any location in the genome. Additionally, they can edit genes by introducing different types of modifications, including gene disruption, gene addition, and allele editing, even in sites that are impossible to be targeted by morpholino nucleotides (Urnov et al., 2010).

8.3 Bistablity of zebrafish tmt opsins

The observation that zebrafish tmt pigments were successfully reconstituted with both 11-*cis* and all-*trans* retinoids suggests that they are likely to function as bistable pigments. This bistable property is a characteristic of the invertebrate pigments (Hillman et al., 1983) and has also been observed in other vertebrate non-visual opsins including opn4 (Mure et al., 2009), opn5 (Yamashita et al., 2010), and parapinopsin (Koyanagi et al., 2004).

Interestingly, spectrophotometric analysis from Chapter 5 showed that the formation of both forms of stable photoproducts absorbed maximally in the UV region at ~370 nm in their dark state. Detection of UV light is important for freshwater fish like zebrafish, as a high level of short-wavelength light (60%) can transmit through shallow water (as shown in Fig. 6.1A). Therefore, tmt opsins might have a role in regulating signalling pathways that are linked to DNA protection in the cells upon stimulation by UV light. In other pigments, like opn5 and parapinopsin, photoisomerisation of the retinal chromophore between the 11-*cis* and all-*trans* forms generally causes a spectral shift in the photoproducts to give an absorption maximum at completely different wavelengths (Yamashita et al., 2010,

Koyanagi et al., 2004). In the case of opn5, one is a UV light-sensitive form coupled with 11-*cis* retinal whilst the other is a visible light-sensitive form that binds all-*trans* retinal. Interestingly, photoisomerisation of both 11-*cis* and all-*trans* retinals induces conformational changes in tmt pigments that absorb wavelengths in the UV range. To confirm this speculation, the type of chromophore isomer in non-irradiated (dark) and irradiated (specifically at the spectral peak) extracted from purified tmt pigments may be determined by high-performance liquid chromatography (HPLC) (Walker et al., 2008). Retinoids can be extracted from the irradiated and non-irradiated solutions by hydroxylamine treatment and dried. The samples would need to be resuspended in a mobile phase prior to separation through a HPLC column for analysis. Thus far, the similar photochemistry observed for the 11-*cis* and all-*trans* forms suggests that this bistable nature of tmt opsins may lead to constant activation of the pigments under UV irradiation.

Based on the bistable characteristic of zebrafish tmt pigments, the possibility that they may also function as photoisomerases cannot be ignored. These opsins may form a newly discovered chromophore retinal-replacement system outside the retina for regeneration of 11-*cis* from all-*trans* form. If this is true, the finding would be significant as this would mean continuous supply of 11-*cis* retinal for maintenance of the broad photosensitivity observed. The system could be important for restoring photosensitivity of other (monostable) pigments, as exemplified by the rhodopsin-retinochrome system in the squid visual cells (Terakita et al., 1989) and the RGR-visual pigment system of RPE, rods and cones that functions independent of the visual cycle (Maeda et al., 2003). Retinochromes of the squid visual cells act to

convert all-*trans* retinal into 11-*cis* forms, which are then transported by retinal-binding proteins to re-conjugate with the rhodopsins for maintenance of their photosensitivity (Sperling and Hubbard, 1975). Collectively, it may be concluded that zebrafish tmt opsins form bistable pigments which may have both G protein activation as well as photoisomerase functions, which has been speculated for invertebrate opsins and other non-visual opsins such as melanopsin and opn5.

8.4 Comparison between zebrafish and cavefish tmt opsins

Both the surface- and cave-forms of *Astyanax mexicanus* exhibit circadian clock gene expression in their peripheral tissues (Beale et al., unpublished data), similar to those observed in the zebrafish (Whitmore et al., 1998). Despite being adapted to living in dark cave environments, with reduced eyes and loss of pigmentation, the circadian clocks of troglomorphic *A. mexicanus* cavefish have been shown to be photoentrainable (Beale et al., unpublished data). These findings from the cavefish together with the discovery of zebrafish tmt opsins suggest that the *A. mexicanus* may also possess tmt opsins that may function as peripheral photopigments. The existence of *A. mexicanus* cavefish provides a natural experiment that may give a clue to the functional link to the opsins. It has been suggested that long-term removal of light from the environment can lead to mutations in Somalian (*P. andruzzii*) *tmt* opsin and *opn4m-2*, causing them to loss their light detection function (Cavallari et al., 2011). However, evidence collected from this study so far suggested that this is not the case; the *A. mexicanus tmt*6 isolated did not have any deletion or insertion mutation. Studies have shown that the Somalian cavefish were isolated from their

surface form about 1 million years earlier than the *A. mexicanus* cavefish (Jeffery, 2009, Colli et al., 2009). Thus, it could be the *A. mexicanus tmt* opsins and *opn4* have not been isolated in the dark long enough to accumulate dysfunctional mutations in the genes.

Homologues of the opsin genes were isolated from embryonic cell lines of *A. mexicanus* cavefish (Pachón and Steinhardt) and surface populations of fish used in this study. As shown in Chapter 7, full-length sequences of tmt6 were obtained from all three populations of *A. mexicanus*, whilst partial sequences of *tmt9*, *tmt 14* and *tmt24* were confirmed. It is possible that there are other existing members of *tmt* opsins in the *A. mexicanus* genome, which can be identified by *in silico* analysis of RNA sequence data since whole genome sequencing of the species has just been completed recently under BioProject: PRJNA89115 at National Centre for Biotechnology Information (NCBI) GenBank. Like zebrafish tmt opsins, the *A. mexicanus* tmt opsin proteins deduced from present data possess all the conserved residues known to be essential for the formation of functional photopigments (see Fig. 7.2). However, spectrophotometric analysis and electrophysiological data indicated that the tmt6 of Pachón cavefish failed to regenerate *in vitro* with both 11-*cis* and all-*trans* retinal chromophores. Explanations for this were proposed in Section 7.3, which include the possibility of different retinal usage, the requirement of cofactors that might not be present in the heterologous expression systems, or they could not function as UV-absorbing photostable pigments. It is more likely that the *in vitro* system used was not sensitive enough to work with these photopigments, as with opn4 and most of the other zebrafish tmt pigments. To fully characterise the

role of tmt opsins in cavefish, it warrants further studies to investigate their retinal binding capability by methods like HPLC and to test with various other G-protein signalling assays (e.g. bioluminescent reporter assay). Further experiments, such as gene knock-out or transgenesis in animals, will also be required to test the function of the tmt opsins in the *A. mexicanus*.

8.5 Conclusion

In summary, the work presented in this project has answered most of the research questions and fulfilled the objectives initially set out. Molecular properties of all six zebrafish tmt opsins were analysed and their distribution throughout early development and in the adult zebrafish retina and brain were studied. The isolation of this group of UV-sensitive opsins, in conjunction with their functional studies provided a better understanding of their putative roles in the zebrafish. Although tmt pigments from other vertebrates (e.g. *Xenopus* and chicken) are not examined yet, it is possible that they can also be sensitive to UV light. Interestingly, phylogenetic analyses showed that tmt opsins have been found in most species except for the eutherian mammals, and the same has been observed for DNA photolyase enzyme (Kuraku and Kuratani, 2011, Lucas-Lledo and Lynch, 2009). Since both genes were lost at the same time, this suggests that perhaps tmt opsins and photolyase are functionally linked. At present, all other UVS pigments reported so far (e.g. sws1 and parapinopsin) are generally expressed in retina and brain. The six UVS tmt pigments discovered from zebrafish were found to be present in nearly all of their tissues, highlighting the importance of broad UV reception in teleost. Although this study

has shown that some of the *tmt* opsins are coupled to G_i-signalling pathway, their physiological functions have not yet been explored. Based on their expression patterns, spectral characteristics and G_i-coupled signalling pathway, it can be postulated that tmt opsins might be associated with circadian photoentrainment, as well as detection of UV-light for a variety of survival functions, including cell repair and protection. Analyses of the effect of tmt opsin knock-outs on the photoperiodic responses or UV-light induced survival functions should be addressed in future studies.

Appendices

Appendix

A.1 Complete cDNA sequences of zebrafish *tmt* opsins

```
Dr_TMT_Ch2L   ATGATTGTGT CCAACTTGAG TGTGCTCAGT TGCAGGAGGA ACAGCGCGCT 50
Dr_TMT_Ch2S   ATGATTGTGT CCAACTTGAG TGTGCTCAGT TGCAGGAGGA ACAGCGCGCT 50
Dr_TMT_Ch6    ATGTTTCCTG AAGAAACTAA TATGAGTTAC ATTTCGAACG GCACGGATGA 50
Dr_TMT_Ch9    ATGTTTTTCG AGCAGGCCGA TTTAAACTAC AGCTTCAACA TGAGCGAGGA 50
Dr_TMT_Ch10   ~~~ATGGTCA CTGTCCCCTT CCTCCGGGAC TCCTCGGTGA ACGGCAGCCT 47
Dr_TMT_Ch14   ATGGTCGTCT ACATCTGGAG TTTGAACATC AGCAGCAAAG ACACTTCTGC 50
Dr_TMT_Ch24   ATGATTGAGT CTAACGTGAG T~~~~~~~~~ ~~~~~~CGGA GTTGCGAGTG 35

Dr_TMT_Ch2L   GTGTCTCGGT ~~~GCAGTGG AGGGA~~~~~ ~~~~CATTTG GAGGCGTCCT 88
Dr_TMT_Ch2S   GTGTCTCGGT ~~~GCAGTGG AGGGA~~~~~ ~~~~CATTTG GAGGCGTCCT 88
Dr_TMT_Ch6    TGACTTG~~~ ~~~~~~~~~C TGTCTGCACT T~~~GAAGAC TGGAGCGACA 85
Dr_TMT_Ch9    AGACCGC~~~ ~~~~~~~~~T TAACTCTTCT GGACGAGGAC TGGAGCGACT 88
Dr_TMT_Ch10   GTTG~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~GACTCTCTC TCTCCCGCCG 70
Dr_TMT_Ch14   GCTTAACCAA AGTGGCAATG TGAGTTCAGG GGATCCTCTG GAGCCCCATG 100
Dr_TMT_Ch24   GTGCGCTGGA ~~~GGAGGTG AAGGGACCGG AGCGCACTTG GATGAAAATC 82

Dr_TMT_Ch2L   CCTCCTATCG CACCCTCAGT CCGACTGGAC ACATCCTCGT GGCGGTGAGC 138
Dr_TMT_Ch2S   CCTCCTATCG CACCCTCAGT CCGACTGGAC ACATCCTCGT GGCGGTGAGC 138
Dr_TMT_Ch6    CTCCTGCTGA GAAACTGTCA CGGACGGGAC ACAATGTAGT CGCTGTTATT 135
Dr_TMT_Ch9    CGCCAATGGA GACGCTCTCA CGCGCGGGCT TCATCGCGCT CTCTGTGTTC 138
Dr_TMT_Ch10   ATCAGACC~~ ~GGATTCAGC CGCGCGGGAT ACACTGTGGT CGCGGTGATT 117
Dr_TMT_Ch14   ACAGCCCACC CGGCCTGAGC AGGACCGGCC ACACGGTGAC CGCCGTCTGC 150
Dr_TMT_Ch24   ACTCAGACCA CAGTCTGAGC CCGACAGGAC ACTTGGTAGT CGCGGTGTGC 132

Dr_TMT_Ch2L   TTGGGATTCA TCGGGACTTT CGGATTCCTC AACAACCTCC TGGTCCTCGT 188
Dr_TMT_Ch2S   TTGGGATTCA TCGGGACTTT CGGATTCCTC AACAACCTCC TGGTCCTCGT 188
Dr_TMT_Ch6    TTGGGATCTA TTCTAATATT CGGGACCCTG AACAATCTTG TAGTCCTCGT 185
Dr_TMT_Ch9    CTGGGATTCA TCATGACTTT CGGCTTTTTT AATAACCTCG TAGTATTAGT 188
Dr_TMT_Ch10   TTGGGAATAA TCTTTGTTTT CGGTTTCCTT TGTAACTTTG TCGTTCTTTT 167
Dr_TMT_Ch14   CTCGGAGCCA TCCTGCTGTT GGGATGCCTG AACAACCTCT TTGTCCTGCT 200
Dr_TMT_Ch24   CTTGGATTCA TCGGAACTTT CGGCTTTCTC AACAACACGC TCGTGCTCGT 182

Dr_TMT_Ch2L   GCTCTTCGGC CGATACAAGG TGCTGCGCTC CCCCATCAAC TTTCTGCTGG 238
Dr_TMT_Ch2S   GCTCTTCGGC CGATACAAGG TGCTGCGCTC CCCCATCAAC TTTCTGCTGG 238
Dr_TMT_Ch6    TCTGTTTTGT AAATTCAAGA CACTGCGGAC CCCCGTGAAC ATGCTTCTGC 235
Dr_TMT_Ch9    GCTTTTCTGT AAGTTTAAGA CGCTGAGGAC GCCGGTTAAC ATGCTGCTTC 238
Dr_TMT_Ch10   GGTGTTCGCT CGTTTTCATG TACTGAGGAC CCCGATTAAC TTGATACTGT 217
Dr_TMT_Ch14   CGTCTTCGCG AGGTTTCGCA CACTGTGGAC TCCTATAAAC CTGATCCTGC 250
Dr_TMT_Ch24   CCTGTTCTGT CGCTACAAAG TGCTGCGCTC GCCTATGAAC TGTCTCCTGA 232

Dr_TMT_Ch2L   TGAACATCTG CTTGAGCGAT CTCCTGGTTT GCGTTCTGGG GACGCCGTTC 288
Dr_TMT_Ch2S   TGAACATCTG CTTGAGCGAT CTCCTGGTTT GCGTTCTGGG GACGCCGTTC 288
Dr_TMT_Ch6    TGAACATCAG CGTGAGTGAC ATGCTCGTGT GTCTATTCGG CACCACGCTC 285
Dr_TMT_Ch9    TTAACATCAG CATCAGTGAC ATGTTGGTGT GTATGTTCGG GACGACCCTG 288
Dr_TMT_Ch10   TGAATATCAT TGTGAGCGAT ATGCTGGTTT GTCTGTTCGG TACACCTTTA 267
Dr_TMT_Ch14   TGAACATCAG CGTCAGTGAC ATACTGGTCT GCTTGTTTGG GACTCCCTTT 300
Dr_TMT_Ch24   TCAGCATCTC AGTGAGCGAT CTGCTGGTCT GCGTACTCGG CACTCCGTTC 282
```

```
Dr_TMT_Ch2L   AGCTTCGCGG CGAGCACGCA GGGGCGATGG CTCATCGGGG ACACTGGATG 338
Dr_TMT_Ch2S   AGCTTCGCGG CGAGCACGCA GGGGCGATGG CTCATCGGGG ACACTGGATG 338
Dr_TMT_Ch6    AGCTTCGCGG CCAGTATCCG CGGCCGGTGG CTGGTGGGGA GACACGGCTG 335
Dr_TMT_Ch9    AGCTTCGCGT CCAGTGTGCG GGGAAGGTGG CTGCTCGGGC GACACGGCTG 338
Dr_TMT_Ch10   AGTTTCGCCG CGTCTGTTCA CGGCCGGTGG CTGACGGGAG TTCACGGATG 317
Dr_TMT_Ch14   AGCTTCGCGT CCAGTCTCTA TGGGAAATGG CTTTTGGGAC ATCACGGCTG 350
Dr_TMT_Ch24   AGCTTCGCCG CCAGCACGCA GGGTCGCTGG CTCATCGGGC GCGCGGGATG 332

Dr_TMT_Ch2L   CGTGTGGTAC GGCTTTGCCA ACTCGCTGTT AGGCATTGTA TCGCTCATCT 388
Dr_TMT_Ch2S   CGTGTGGTAC GGCTTTGCCA ACTCGCTGTT AGGCATTGTA TCGCTCATCT 388
Dr_TMT_Ch6    CATGTGGTAC GGCTTCGTCA ACTCATGCTT TGGGATTGTA TCGCTGATTT 385
Dr_TMT_Ch9    CATGTGGTAC GGCTTCATCA ACTCCTGCTT CGGTATTGTT TCTCTCATAT 388
Dr_TMT_Ch10   CAGGTGGTAT GGTTTTGCCA ACGCGTTGTT CGGTATCGTT TCACTGGTAT 367
Dr_TMT_Ch14   CAAATGGTAC GGTTTCGCCA ATTCGCTTTT TGGAATTGTG TCTCTGATGT 400
Dr_TMT_Ch24   CGTGTGGTAC GGTTTCATCA ATTCATTCCT GGGCGTTGTG TCTCTAATCT 382

Dr_TMT_Ch2L   CTCTGGCTGT GCTGTCATAC GAACGCTACT GTACAATGAT GGGCTCCACA 438
Dr_TMT_Ch2S   CTCTGGCTGT GCTGTCATAC GAACGCTACT GTACAATGAT GGGCTCCACA 438
Dr_TMT_Ch6    CTCTGGCAAT CCTTTCGTAT GATCGTTACA GTACATTAAC TGTTTACAAC 435
Dr_TMT_Ch9    CCTTGGTGGT TCTTTCTTAC GACCGCTACA GCACATTGAC TGTGTACCAT 438
Dr_TMT_Ch10   CTCTTGCTGT GCTGTCATAC GAGCGTTATA GTACCATCCT CTGCTCCTCA 417
Dr_TMT_Ch14   CTCTTTCTAT ACTGTCATAC GAGCGTTATG CTGCCCTGCT GCGGGCCACC 450
Dr_TMT_Ch24   CTCTGGCGGT TCTCTCATAC GAGCGTTACT GCACCATGAT GGGCTCCACA 432

Dr_TMT_Ch2L   GAGGCTGATG CGACCAACTA CAAGAAGGTG ATCGGTGGGG TGCTGATGTC 488
Dr_TMT_Ch2S   GAGGCTGATG CGACCAACTA CAAGAAGGTG ATCGGTGGGG TGCTGATGTC 488
Dr_TMT_Ch6    AAGAGG~~~G CTCCAGACTA CAGTAAACCC TTGTTGGCAG TCGGGGGCTC 482
Dr_TMT_Ch9    AAACGG~~~G CTCCAGACTA CCGCAAGCCT CTGCTGGCAG TAGGAGGCTC 485
Dr_TMT_Ch10   AAAGCAGATG CATCTGACTA TAGGAAGGCC TGGCTCTTCA TCACAGGCTG 467
Dr_TMT_Ch14   AAGGCCGATG TGTCTGACTT CCGCAGGGCC TGGCTCTGTG TGGCTGGTTC 500
Dr_TMT_Ch24   CAGGCCGACT CCACCAACTA CCGAAAAGTG GTCATAGGCA TTGCTTTCTC 482

Dr_TMT_Ch2L   CTGGATCTAC TCTTTAATTT GGACCCTGCC TCCTCTGTTT GGGTGGAGCC 538
Dr_TMT_Ch2S   CTGGATCTAC TCTTTAATTT GGACCCTGCC TCCTCTGTTT GGGTGGAGCC 538
Dr_TMT_Ch6    GTGGCTCTAC TCACTGTTCT GGACGGTTCC TCCTCTGCTG GGCTGGAGCA 532
Dr_TMT_Ch9    CTGGCTGTAC TCTCTGATCT GGACAGTGCC TCCTCTTTTA GGCTGGAGCA 535
Dr_TMT_Ch10   CTGGCTGTAC TCTCTGCTAT GGACTGTACC ACCGCTCTTA GGCTGGAGCA 517
Dr_TMT_Ch14   CTGGCTATAT TCACTGCTAT GGACTCTTCC ACCATTTCTG GGATGGAGCA 550
Dr_TMT_Ch24   CTGGATTTAT TCCATGGTTT GGACTCTTCC ACCACTGTTC GGATGGAGCT 532

Dr_TMT_Ch2L   GCTATGGCCC CGAAGGCCCC GGGACAACCT GCTCTGTAGA CTGGACGACC 588
Dr_TMT_Ch2S   GCTATGGCCC CGAAGGCCCC GGGACAACCT GCTCTGTAGA CTGGACGACC 588
Dr_TMT_Ch6    GCTATGGACT GGAAGGGGCA GGAACCAGCT GTTCGGTGAC ATGGACTGCC 582
Dr_TMT_Ch9    GTTATGGGCT GGAGGGTGCC GGAACCAGCT GCTCTGTATC ATGGACTCAG 585
Dr_TMT_Ch10   GCTATGGTCC TGAGGGTCCC GGCACCACCT GTTCAGTCCA GTGGAACAAG 567
Dr_TMT_Ch14   ACTATGGTCC TGAAGGACCC GGGACTACAT GCTCAGTGCA ATGGCACCTG 600
Dr_TMT_Ch24   GCTACGGTCC GGAGGGTCCC GGCACCACCT GCAGCGTCAA CTGGGCCGCG 582

Dr_TMT_Ch2L   AAGACCGCCA ATAACATCTC CTACATCATT TGCCTCTTTA TTTTTTGCCT 638
Dr_TMT_Ch2S   AAGACCGCCA ATAACATCTC CTACATCATT TGCCTCTTTA TTTTTTGCCT 638
Dr_TMT_Ch6    AACACGCCCC AGTCGCATTC CTACATCATC TGCCTGTTCA TTTTTTGCTT 632
Dr_TMT_Ch9    CGCACGGCCG AGTCTCATGC CTACATCATC TGTCTGTTTG TATTCTGCCT 635
Dr_TMT_Ch10   CGCTCTCCGG AAACCCGGTC TTACGTCATC TGCCTCTTTG TCTTCTGTCT 617
Dr_TMT_Ch14   CGCTCAACCA GCAGCATCTC GTACGTCATG TGCTTGTTCA TCTTCTGTCT 650
Dr_TMT_Ch24   CGGACTCCCA ACAACGTGTC TTACATTGTC TGTCTGTTTG TCTTCTGCCT 632
```

```
Dr_TMT_Ch2L   GATCGTCCCC TTCCTGGTCA TCATATTCTG CTATGGGAAG CTGCTGCATG 688
Dr_TMT_Ch2S   GATCGTCCCC TTCCTGGTCA TCATATTCTG CTATGGGAAG CTGCTGCATG 688
Dr_TMT_Ch6    AGGGATCCCT GTGCTGGTCA TGGTGTACTG TTACAGTCGG CTCATCTGTG 682
Dr_TMT_Ch9    GGGACTACCA GTGCTTGTCA TGGTCTACTG CTATGGGAGA CTGCTGTACG 685
Dr_TMT_Ch10   TCTGCTGCCT CTCCTTCTGA TGGTCTACTG CTATGGCAAG ATCCTCATTG 667
Dr_TMT_Ch14   GCTTCTGCCC CTCGTGCTCA TGATCTTCTG CTATGGCAAA ATCCTGCTAC 700
Dr_TMT_Ch24   TATTCTCCCT TTTATCGTGA TCGTCTACAG CTACGGCAGA CTCCTTCAGG 682

Dr_TMT_Ch2L   CCATTAAACA GGTCAGCAGT GTGAACACGT CTGTGAGCCG CAAGCGTGAG 738
Dr_TMT_Ch2S   CCATTAAACA GGTCAGCAGT GTGAACACGT CTGTGAGCCG CAAGCGTGAG 738
Dr_TMT_Ch6    CTGTCAAACA GGTAGGCCGC ATTAGGAAGA CAGCAGCAAG ACGGAGAGAG 732
Dr_TMT_Ch9    CCGTCAAACA GGTAGGGAAG ATACGCAAGA CGGCTGCACG GAAACGGGAG 735
Dr_TMT_Ch10   CCATTCATGG GGTTGCTAAG ATCAATCAGA CAGCTGCACA GCGGAGGGAA 717
Dr_TMT_Ch14   TCATCAAAGG GGTAACAAAA ATCAACCTGC TGACTGCGCA GAGGAGAGAG 750
Dr_TMT_Ch24   CTATCACACA GGTGAGCAGG ATCAATACGG TGGTGAGTCG TAAACGGGAG 732

Dr_TMT_Ch2L   CACCGTGTCC TACTGATGGT CATCACTATG GTGGTTTTCT ATCTGCTGTG 788
Dr_TMT_Ch2S   CACCGTGTCC TACTGATGGT CATCACTATG GTGGTTTTCT ATCTGCTGTG 788
Dr_TMT_Ch6    TACCACATTT TATTTATGGT CATCACTACA GTGGTGTGCT ACCTTTTGTG 782
Dr_TMT_Ch9    TACCATGTGT TGTTCATGGT TATCACCACT GTGGTGTGTT ACCTTTTGTG 785
Dr_TMT_Ch10   ACACACGTGC TGGTGATGGT GGTTTCCATG GTGTCCTGTT ACCTGCTCTG 767
Dr_TMT_Ch14   AACCACATTC TTCTGATGGT TGTCACTATG GTATCCTGCT ATCTGCTGTG 800
Dr_TMT_Ch24   CAGCGTGTGC TCTTCATGGT GGTCACCATG GTGGTCTGTT ACCTGCTTTG 782

Dr_TMT_Ch2L   CTGGCTCCCC TACGGCATCA TGGCCCTGCT GGCCACTTTT GGAGCTCCAG 838
Dr_TMT_Ch2S   CTGGCTCCCC TACGGCATCA TGGCCCTGCT GGCCACTTTT GGAGCTCCAG 838
Dr_TMT_Ch6    CTGGATGCCT TATGGAGTTG TTGCAATGAT GGCCACGTTC GGACGCCCAG 832
Dr_TMT_Ch9    CTGGATGCCG TACGGCGTCG TGGCCATGAT GGCCACATTT GGCCGGCCGG 835
Dr_TMT_Ch10   CTGGATGCCG TACGGGGTCA TGGCATTGCT CGGCACATTC AGC~~~GCTG 814
Dr_TMT_Ch14   CTGGATGCCT TACGGGGTGG TGGCCCTGCT GGCCACCTTT GGCAGAACGG 850
Dr_TMT_Ch24   CTGGTTGCCA TATGGGATAA TGGCGCTGCT GGCTACGTTT GGACATCCCG 832

Dr_TMT_Ch2L   GACTCGTGAC GGCGGAAGCC AGCATAGTGC CCTCCATCTT AGCCAAGTCC 888
Dr_TMT_Ch2S   GACTCGTGAC GGCGGAAGCC AGCATAGTGC CCTCCATCTT AGCCAAGTCC 888
Dr_TMT_Ch6    GAATCATCTC ACCCATTGCA AGCGTGGTGC CATCCCTCCT TGCCAAAAGC 882
Dr_TMT_Ch9    GAATCATCTC ACCTGTGGCC AGTGTGGTTC CCTCTCTCCT CGCCAAGAGC 885
Dr_TMT_Ch10   GCATTACCAG CCCCACAGCC AGCGTCGTGT CATCCTTACT TGCAAAGAGT 864
Dr_TMT_Ch14   GACTAATCAC TCCTGTAACC AGCATAGTGC CATCAGTCCT GGCCAAGAGC 900
Dr_TMT_Ch24   GACTGGTGAC ACCAGCAGCC AGCATAGTTC CTTCTCTGTT AGCCAAGAGC 882

Dr_TMT_Ch2L   AGCACAGTCA TCAATCCTGT CATCTACATC TTTATGAATA AACAGTTCTA 938
Dr_TMT_Ch2S   AGCACAGTCA TCAATCCTGT CATCTACATC TTTATGAATA AACAGGTGAG 938
Dr_TMT_Ch6    AGCACAGTCA TCAATCCTCT CATTTACATC CTCATGAACA AACAGTTTTA 932
Dr_TMT_Ch9    AGCACGGTCA TCAACCCACT CATATACATC CTCATGAACA AACAGTTCTA 935
Dr_TMT_Ch10   AGCACCGTTC TGAACCCCAT CATCTATGTA TTGTTCAATA ATCAGTTTTA 914
Dr_TMT_Ch14   AGCACTGTGG TCAACCCAGT AATCTATGTG CTTTTCAACA ACCAGTTCTA 950
Dr_TMT_Ch24   AGCACCGTCA TAAACCCAAT CATCTACATC TTCATGAACA AACAGTTTTG 932

Dr_TMT_Ch2L   CAGGTGTTTC AGAGCTCTTC TCAACTGCGA CAAACCACAA CGAGGTTCTA 988
Dr_TMT_Ch2S   C~~~ATGTGG AAAGATATAG CTGGATTTTG A~~~~~~~~~ ~~~~~~~~~~ 966
Dr_TMT_Ch6    CAGGTGTTTC CTCATCTTGA TTCACTGCAA ACATAGCTCT CTAGAGAACG 982
Dr_TMT_Ch9    CAGGTGTTTC CGCATCTTGT TCTGCTGTCA GAGAAGTTTG CTACAGAACG 985
Dr_TMT_Ch10   CAGGTGCTTT ATAGCATTAG TAAGAAGCGG AGCAGAACCT CCAGTTCAT~ 963
Dr_TMT_Ch14   CAGGTGTTTC GTAGCATTTC TGAAATGTCA AGGCGAACCA TCTGTCCATG 1000
Dr_TMT_Ch24   CAGATGTTTC CATGCGCTCA TCATGTGCAC CACCCCAGAG AGAGGCTCCA 982
```

```
Dr_TMT_Ch2L   GTCTGAAAAG CTCTTCAAAG ACC~~~AAAC CTTTTCGTCC CGGACGCCGC 1035
Dr_TMT_Ch2S   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Dr_TMT_Ch6    GACAGTCATC AATGCCTTCA AGAACTACAG GCATCCAGCT TAATCGAAGG 1032
Dr_TMT_Ch9    GTCACTCCTC CATGCCCTCC AAGACCACCG TCATCCAGCT GAACCGCAGG 1035
Dr_TMT_Ch10   ~~~~~~~~~~ ~CTCACCCTT CACACAGAAG AAGGTGCTGC TCAACAACAC 1002
Dr_TMT_Ch14   GACAGAATCC ACAGCATAGC AGCAAAGAAG ATCCACACGT GTTCAGGCCC 1050
Dr_TMT_Ch24   GCTTTAAAAA CTCTTCCAAA GTCACCAAAA CTCTCAGGAC AGTGCGACGT 1032

Dr_TMT_Ch2L   ACGGAC~~~~ ~~AACTTCAC TTTTATGGTT GCTTCAGTGG GACCC~~~~~ 1074
Dr_TMT_Ch2S   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Dr_TMT_Ch6    CCTTACAGCA ACCCAGTGGC TGACAACGCA CCC~~~~~~~ ~~~~~~~~~~ 1065
Dr_TMT_Ch9    GTCAACAGTA ACGCTGTAGC CTGCACCGCT CAG~~~~~~~ ~~~~~~~~~~ 1068
Dr_TMT_Ch10   TGCATCCCTA TGGGATTATA TGCTGCTACC TCT~~~~~~~ ~~~~~~~~~~ 1035
Dr_TMT_Ch14   TGTGACGGGG CCTCAATTCA TCGCAGCGCA GAG~~~~~~~ ~~~~~~~~~~ 1083
Dr_TMT_Ch24   GCCAACGGCC AGAATGTGAC CTTCGCTGTG GCATCGGCCG TCCATCGAAC 1082

Dr_TMT_Ch2L   ~~~~~~~~~~ ~~~~~~~~~A ATCAGACAAA CCCAGTGGAG GACGGCCCAC 1105
Dr_TMT_Ch2S   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Dr_TMT_Ch6    ~~~~~~~~~~ ~~~~~~~~~~ ~~CCTTCCAT CGATCTCCAA AAC~~~GATT 1090
Dr_TMT_Ch9    ~~~~~~~~~~ ~~~~~~~~~~ ~~ATCTCCAC CGGGACTCAC AACCATGATT 1096
Dr_TMT_Ch10   ~~~~~~~~~~ ~~~~~~~~~~ ~~CCACCTGA G~~~~~~~~~ ~~~~~~AGCC 1048
Dr_TMT_Ch14   ~~~~~~~~~~ ~~~~~~~~~~ ~~GGCCCTCA GAAAAAAGAG CAGCACTCTC 1111
Dr_TMT_Ch24   TCCGTACAGC GACAGACAGA AAAGTTCCTC GGAGGGTGAG AAACTCCCTC 1132

Dr_TMT_Ch2L   CATCAGCAGA TAAC~~~ACC AAACCTGCAG TGCTCTCCCT CGTGGCCCAT 1152
Dr_TMT_Ch2S   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Dr_TMT_Ch6    GCAGCACCCC TGTCTCTGGC TAA~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ 1113
Dr_TMT_Ch9    GCAGCACTCA CGTCACAGAG AGGAGCAACC CTCCAGAGGT CATTCCATGA 1146
Dr_TMT_Ch10   CAAGCCTGAT GGACACCCCC AAATGA~~~~ ~~~~~~~~~~ ~~~~~~~~~~ 1074
Dr_TMT_Ch14   TGTCCTTGGT GGTTCACTAC ACACCCTGA~ ~~~~~~~~~~ ~~~~~~~~~~ 1140
Dr_TMT_Ch24   CAGCTACGGG TCAGGGCACC AGCAAACCTG TTGTCTCTCT TGTGGCCTAT 1182

Dr_TMT_Ch2L   TATAACGGAT GA 1164
Dr_TMT_Ch2S   ~~~~~~~~~~ ~~
Dr_TMT_Ch6    ~~~~~~~~~~ ~~
Dr_TMT_Ch9    ~~~~~~~~~~ ~~
Dr_TMT_Ch10   ~~~~~~~~~~ ~~
Dr_TMT_Ch14   ~~~~~~~~~~ ~~
Dr_TMT_Ch24   TACAATGGCT GA 1194
```

A.2 5′-promoter analyses of *tmt2* and *tmt6*

(A) 5′-promoter region of *tmt2*

TTGTAAAACCTGGCAACGCTGAGTAACAGTGAAGCGAACAGACCACAAGTACGTTCCTAACCACAGTCGACGCTT

nr2e3

TGGAAAAATCCTGAACTGCTCAAAAGTCAAAGTTTTTTTGTTATGGAGCAAACTTAAGCACCATGGACAGCGCTC

AGACCAAATATGTGAACAAAAATAAGGGATTCAATTTAAGTTATGGATGTCAAACTTGCCCTTGTCTTAAATTGA

ACATTGCCCTCGAGAGAGATAGAGAGTAAACTAAAATAAATGTTTTGAATTTTTTTCCTCAGGCTTCTATGTGTA

d-box

CAGTAAGTTTATTAACTTTCTTTTATCATAAATAATAGGTACTTTTCATTGCCTTTAAATTGAGTTACTGAACCT

ret-1/pce-1

ACACACACAAACACAAACTAACCCGAGAAAAAAAAGATCAATTACAAAAAAAGAATTGGAATCTGAGAGTCCAAA

crx

AACAAAATAGAAATACTGAGGAAAATGGCTAATGTTGTCTAGTGCTCAGTAATACACATTACTAATCAAACAACG

HNF-1

GGTTTTGATATATTAACAGAATGACATTTACTTAATACCTACATAGAACATGACTATGGTCTTTACTTAAAATCC

ror-β

AAATGAAATTTGGCATAACAAAATCTTTCATTTTGACCCATAAAATGTATTTTTTTGGATATTACCAAAAATTAC

ret-4 crx

CTGTGCAGAATAAGTCTTTGGGTTTATGGTCCAGGGTCACATTTGCTTACATAAATCACATGAAGTGAGGGCGCT

HNF-4

GTTCATGGACTGGACTGGACTTAAGTTGAGTTTTGATTGGCCTTTGGGCAACATTFGTGGTGTGAAAACCTGGCA

ACCCTGAGTAGCAGTAAAGCCAACAGAACACAAGTACGTTCCCATACAGTCAACGCTTTGAAAAAATCCGGAACT

GCTCAAAAGTCGATGATTTTTTTCTTATGGAGCAAACTTAAGCACCATGGACAGCGCTCAGACCAAACGTGAACA

AAAATAAGGGATTCAATTTAAGTTATGGATGTCAAACTAGCCCTTGTCTTTCACTGAACATTCCCCTCGAGAGAG

ATTGAGAGTTAAACTAAAATAAATGTTTTGCACTTTTTTTCTCAGGCTTCTATGTGTACGTTCAGTAACTTTCCT

TTATCATAAATCTTATGTTCTTTTTATTGCCTTTAAAGTGAGTTAATGAACCTACACACACAAACACAAACTAAC

CCGAGAAAAAAAAGGATGATTTAAAAAAGAATTGGAATCTGAGAGTCCAAAAACAAAATAGAAATACTGAGGAAA

crx otx/bat-1 crx

ATGGCTTTAATCTTGTCTAGTGATTAGTAATGCACATTACTAATCGAACAACGGGTTTATTCACAGCAGGACATT

TACTTAATACCTTCATAGAACATGACTGTGATCTTTACTTAAAATCCTGATGAAATTTTGCATTACAAAATCTTT

chx 10

ATTTAATTTTGACCCAAAACAAATGTAATTTTGAACATAACCAAAAGCATAAGTGTGCAGAATAATTCTTTTGCC

ret-4 e-box and pax-6

TTGTGGTCCAGGGTCACATTTGCTTATCTTGCGATTACAATAACACGTGAAGTGAGGTGTTTGTTTTTTGTGTTA

ret-4

GTTTAAATGGACATTGGGCTAGTTTTAACTGACCATTTCCCTGCTGAAACATCCAGCTTAAACCAGCCTAGGCTG

GTTTAAGCTGGATTTTTCACCAGGGTTGGCTACCTTTGTTGTGAAAACCTGGCAATCCTGAGTAACAGCGAAGCA

HNF-3β

AACAGAACGCAAGCATGTTCCCATGCAGTCTACGCTTTAAAAAAACACTGAACTGCTCAAAAAGTTGATGTTTTT

TGTTTCTCACGGAGCAAACTTTGAGCACCACAGCGCTCAGACCAAACATGTGAACAAACATGAGCAATTCAATTT

AAGTTATCGATGTCAAACTTTTCCTTGTCTTAAAAAAACATCTTGAACACAAACATTGGGTTTTAATATATTTAC

ret-4 e-box

AGTAGGACATTTACTTAATATCCAGGGTCACATTTGCTTATCACATGATTACAAAAAACACGTTAAGTGAATTGA

HNF-4

GGTAAAGTTGGTTTTATATTCACACATTTAAACTTTGTCCTTTATAATTTCAGAAAGGTATTATTTGCAAATTTT

ret-1/pce-1

AAGTAGGGAAATGATATTAGCAGCCAATATCAAAGCACAGTCAATTAAATGTGTTAATGTTGTTTAGAAGTCATA

chx-10

CTTGCATGTTAATTCAATATTAAATGTACAAAACAGAGACTGGGTCACAAGTTGTCATTGTTTAAAAATGAATTA

ACGTGCGAACATAAAACCAACTTTACCTCAATGTGAAGTGTGAGGGCACTGTTCAGGGACTGGGCTTGAATTTGG

GCTAGTTTGAAAGGACACTGGGGTACTTTTGTTGTAAAACCTGGCAACCCTGAGTAACAGTGCAGCGACCAGACC

ACAAGTACGTTCCTTATCACAGTCGACGCTTTGGAAAAACCCTGAACTGCTCAAAAGTCGATGTTTTTTCCCCTC

pax-4

GTGGAGCAAACTTGAGCACCATGGACAGCGATCAGGCGTGCGAACAAAAATAAGCGATTCAAATGAAGTTATGAA

TGTCACACTGACTTCCCCATCTCTGAATATTCCCCTCGAGAGCGAGAAAATTAATGTATTGCAAACCATTTATTG

nrl

CGTCGTGCGTCAAAACACAGACGGGCAGATGTGCTGCAGAGGTTTTCTGCTCATAACCTGGATGATGCGTCCAGT

GC-box

AAAGCTCAGTAGAGCAGACAGACTCGGGCGTCAGAGAGGAGGTACACTACTTACACCAACAGTACACTTTCACCG

BRE

CACATCAGGACTGTAAACACCACACGGTGGAGCAAAGGGACGCCAAACTTTATTCCTACTTTTAAAGTTGAGCAA

TATA-box Initiator element DPE

TTGCTAAATAATCGTCGCGCAACACCCAAACTCTTCAGCAAAAATACGATTAGATTTGCAGACTGTTGTCAAAAC

AP-1

AGCACATCAAATGTGTCGCTTTCCCCTCAGAAAACAAAGCAGTGACTTATTGTGCTCAATGGATTGCGGATTGCG

ATG

(B) 5′-promoter region of *tmt6*

TGTTCCACTAGATCCCAAAGGTGCTTTATTGAGTTTTTATTGAGTTTTGGTGACTGTGGAGGCTATTTGAGTACA

rorβ

GAAAACTCATTGTCGTCTTCCTGAAACCAGTCTGAGATGATGCCATCGGAAGATGGGTAAACTGTGGTCATAAAG

GGATGGACATGGTCGACATCGATACTCAGGCTGTGGTGTTGGCACGATGCTAATGGTACTAATGGGCCCAAAGTT

TGCCAAGAAAATATCCCCTACATCATTACTCCACCACCACCAGCCTGAACCATTGATACAAAGCAGAATGGATCC

chx-10 glass-like binding motif

ATGCTTTAATTTTGTTGACACCAAATTCGGACCCTACCCTCCAAATGTGGCAGCAGAAATGGAGATTCAGAACAG

GCAACATTTTTTCAATTCTCTTTTGTCCAGTTTTGGTGAGCCTGTGTGAATTGTAGCCTACGTTTCCTGTTCTTA

CCTAACAAGAGTAGCACCCAATGTGGCATTCTGCTTCTGTAGCCCATCTGCCTCAAGTTTCGGCATGTTGTGTGT

otx/bat-1

TCAGAGATGCTTGTCTTCATACCTCGCTTGTAACAAGTGATTATTTGAGTTACTCTTGTCTTTCTATCAGCTCGA

rorβ

CCCAGTCGGGTTATTCTCCTCTGACCTCAGGCATCAACAAGGCATTTGTGCCCACAGAATTGCCGGTCACTAGAT

ATTTTCTTTTTTCAATCCATTCTCTGTAAACCCTAGAGATGGTTGTGCAGGAAAATCCCAGTAGATCAGCAGTTT

ret-4

TTGAAATACTCAGACTAGCCTGTCTGGCACCAACCATGTCAAATTACTTTCTTCCTCATTCTGATGCTTGGTTTG

AACTGCAGCAGATCGTCTTGACTGTCTACATGCCTTAATGCATTGTTGCGGCAAGTGTACCCAATAAAGTGACCA

TTAGGTGTATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA

TATATATATATATATATATATATANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATTATAAGTTTTGTTTGGACATTGCA

chx-10

ACTAATTGGACTTTAGAGTATATTATTAGACTGTTAGTTTAGTGTTAGTGTAAGTTGATGGGCTGGCAAAGTTTA

TTATAGGCAGGAATATTTCTGTTAGTAAAGCAGATATTGGGCAGTATATTATCCTCTAATGAGAATAATTGGTAA

CACTTTACAATAACAGTACATGAATAATGATGTATCAATATGTAAACTAAACATTACTTTATTATGAATTAATGA

pax-4

TGAGTTAAGGTATGCGCTAATTATGAATTAACTTTAATTACAACATGATTTACATGAGTTCATGTGTGAATAATG

HNF-1 pax-6

GCTACCTTAATTACTTGTAAGTTCAGGATTGTTTGTTAATTAATGTATTAATTACCACATTAGTTTATGCTTTAT

d-box

TTAACCATGATGAACTCATGATATGTTAATGTATAATTATGTTATGACTTTACTTGGAGGGGCACATCATCATTA

pax-4

ACCGATTCTTAACTACTCATGAAGTCCTTATGCACGTCTAACTAGTGAGTTTACACTGAATAACAATAGAAAAGT

GTTCTATCAACATCAACATGAAGAGGAGTTCATAAGTAGTTAAGGATGAGTTAATAGTGATGTGCCCCTCCAAGT

nrl HNF-1

AAAGTCATGACATAACTATACATTAACAGCTCATGAGTTCATGTTGGTTACATTTAGCCTAAACTTAAATGTTAA

pax-4 ap-1

TTAATACATTTATTAACAACATGTACTAACAAGTAATTAAGGTAGTCGTTATTTACACATGAACTCATGTGACTC

otx/bat-1

ATGTTGTACTTAAAGTTAATTCATGATTAGTGCATACCTTAACTCATCATTAATTCATAATAATGTTTAGTTTAC

ATATAAATACATCATTATTTATGTACTGTTATTGTAAAGTGTTACCGAATAATTTAGCATAGTTGTTGCAACGTA

crx

ACTTGTATTCTACTAAATGTGCGTCACCATTAAAATGAAGTAACCCCCCCTCCCTCCAAATAATTCAATCAATAA

ATAAAAACAAAATAAAACCTTCAGATAGTTAATTAAAGTGCTAACGTGCAGTTGCTACATTATATTATGTAAAGA

CATAACCCATAATAAGTGTACTGGATGTTTAACAAAATTTTGTGCTAAATAAAAATGCTCTAATTTAGATGTTTC

crx elk-1 oct-1 ret-1/pce-1

ACTTAATCTAATACGTTGACCGGAAATTTATCAATAAAATTAATGCAAATCGAGCAATTATTGTTGTTACAACTA

crx

CAATACTGCAATCATAGTGTCCACCCAAAGCCCTTCCTTGAAAATGTAAAGCCCCGCGCATCATATGGAGTGTCG

nkx2-5 GC-box

ACGTCACTCATTATTATCCTCTCAAGGCTTCTTCCACACGTAAGCGCTCTCAAGTGCTCTGAAGAGGCGGCAAAG

TCGAGGTGCGCGAGCCTTGCAGCGGGGATAAAAAAAAAAACTTGATTTCGTTCTTACAGGCTTCTTAATAACCCT

chx-10

TTTATCTGCACACCTGTCATTTAGTCTATCCAACAGCATGTCAAGGTAATTGATTTTAGTTTGTTTTACTTCTGT

CAGTGTACATTCTGACTGTATGAAGGCTAGCTCACCTGTATTGTTTGATATTCTGTAAAGCTGTTTGCTGTTTAC

ret-4

GTTTACGCGTTGCACCGGCATCCTCTCAATTGAGCAACAACTTAGCTTAACCTGTTTGAAATTTACTAATGGCTT

TATA-box Initiator element pax-6

TATATTTGGGCTGCAAAGCCCTTGCGACACTATTCTTGCGCTGTGAAGACGCTTGTGAAGATTTGACGACGGAAT

TATGAAGCATGAATTGAGACTCTTCAGAGAACTTGATCAACTCTTCACACAGCATGGTCCACTCTCGGTTGTGGA

GTAGCTCCGATAGAATTTTATCCATGGATCAAGGACGCAGTAAACGTAACATCAGTTGAAAACTGTGCGGCAAAG

ATG

Figure A.2 Promoter analyses of 5′ flanking region (3 kb) of zebrafish **(A)** *tmt2* and **(B)** *tmt6*. The main start codon (ATG) of each sequence is bolded. The core and proximal promoter elements that were identified are highlighted in red, which include: an initiator element (INR) with the main transcription start site (A) underlined (Liston and Johnson, 1999); a putative TATA-box (Juo et al., 1996) at 16 bp (*tmt2*) and 20 bp (*tmt6*) immediately upstream of the main transcription start site; a GC box) (Sogawa et al., 1993); B recognition element (BRE) (Lagrange et al., 1998); Downstream Promoter Element (DPE) (Burke et al., 1998); Oct-1 binding site (Phillips and Luisi, 2000); and AP-1 binding sites (Lee et al., 1987). In addition, tissue-specific regulatory sites were also identified, consistent with gene regulation in the development of various tissues. These include the homeobox genes essential for retinogenesis (e.g. pax-4 and pax-6) (Grindley et al., 1995, Kozmik, 2005), factors regulating liver functions (e.g. HNF-1, HNF-3β, HNF-4) (Costa et al., 2003) and lung/testis development (e.g. elk-1) (Rao et al., 1989), genes involved in development and differentiation of photoreceptors (e.g. crx, ret-1/pce-1, otx/bat-1, rorβ, ret-4, glass-like, nr2e3, nrl) (Zhu and Craft, 2000, Swaroop et al., 2010, Akimoto, 2005) and bipolar cell (e.g. chx-10) (Dorval et al., 2006), as well as the sites which are important for circadian clock functions (e.g. d-box and e-box) (Zhang et al., 2004).

A.3 Nucleotide sequences of *Astyanax. mexicanus tmt* opsins

(A) Full-length *tmt6*

```
Dr_tmt6           -------ATG TTTCCTGAAG AAACTAATAT GAGTTACATT TCGAACGGCA 43
Am_SF_tmt6        ATGTTCTATG ATCTGGAGAT CTCCA-ACTT CAGCACTAAC TGGAGCACAG 49
Am_PC_tmt6        ATGTTCTATG ATCTGGAGAT CTCCA-ACTT CAGCACTAAC TGGAGCACAG 49
Am_SH_tmt6        ATGTTCTATG ATCTGGAGAT CTTCA-ACTT CAGCACTAAC TGGAGCACAG 49
Clustal Consensus        *** :*    ..*  .: *: * :*  ** :. *:  * **.*. ..

Dr_tmt6           CGGATGATGA CTTGCTGTCT GCACTTGAAG A--------- ---------- 74
Am_SF_tmt6        GTGAGGAGGA TGCTGAGGGG GAGGAAGGTG CGCTGTCTGC GTTGGGCTGG 99
Am_PC_tmt6        GTGAGGAGGA TGCTGAGGGG GAGGAAGGTG CGCTGTCTGC GTTGGGCTGG 99
Am_SH_tmt6        GTGAGGAGGA TGCTGAGGGG GAGGAAGGTG CGCTGTCTGC GTTGGGCTGG 99
Clustal Consensus   ** ** **      :*    *.. ::*.:* .

Dr_tmt6           --CTGGAGCG ACACTCCTGC TGAGAAACTG TCACGGACGG GACACAATGT 122
Am_SF_tmt6        GGCTGGAGCG ACGCCCCCGA GCAGAGGCTG ACCCGCACCG GGCACAGCGT 149
Am_PC_tmt6        GGCTGGAGCG ACGCCCCCGA GCAGAGGCTG ACCCGCACCG GGCACAGCGT 149
Am_SH_tmt6        GGCTGGAGCG ACGCCCCCGA GCAGAGGCTG ACCCGCACCG GGCACAGCGT 149
Clustal Consensus   ******** **.* ** *.   ***..*** :*.** ** * *.****. **

Dr_tmt6           AGTCGCTGTT ATTTTGGGAT CTATTCTAAT ATTCGGGACC CTGAACAATC 172
Am_SF_tmt6        GGTGGCCGTC ATCCTGGGCT TCATCATGGT GTTCGGCTTC CTGAATAACG 199
Am_PC_tmt6        GGTGGCCGTC ATCCTGGGCT TCATCATGGT GTTCGGCTTC CTGAATAACG 199
Am_SH_tmt6        GGTGGCCGTC ATCCTCGGCT TCATCATGGT GTTCGGCTTC CTGAATAACG 199
Clustal Consensus .** ** **  **  * **.*   ** .*..* .***** : * ***** **

Dr_tmt6           TTGTAGTCCT CGTTCTGTTT TGTAAATTCA AGACACTGCG GACCCCCGTG 222
Am_SF_tmt6        TGGTGGTGCT GGTGCTGTTC TGCAGGTTTA AGACGCTGCG CACCCCCGTT 249
Am_PC_tmt6        TGGTGGTGCT GGTGCTGTTC TGCAGGTTTA AGACGCTGCG CACCCCCGTT 249
Am_SH_tmt6        TGGTGGTGCT GGTGCTGTTC TGCAGGTTTA AGACGCTGCG CACCCCCGTT 249
Clustal Consensus * **.** **  ** *****  ** *..** * ****.*****  ********

Dr_tmt6           AACATGCTTC TGCTGAACAT CAGCGTGAGT GACATGCTCG TGTGTCTATT 272
Am_SF_tmt6        AACCTGCTGC TGCTCAACAT CAGCGTCAGC GACATGCTGG TGTGCGTGTG 299
Am_PC_tmt6        AACCTGCTGC TGCTCAACAT CAGCGTCAGC GACATGCTGG TGTGCGTGTG 299
Am_SH_tmt6        AACCTGCTGC TGCTCAACAT CAGCGTCAGC GACATGCTGG TGTGCGTGTG 299
Clustal Consensus ***.**** * **** ***** ****** **  ******** * ****  *.*

Dr_tmt6           CGGCACCACG CTCAGCTTCG CGGCCAGTAT CCGCGGCCGG TGGCTGGTGG 322
Am_SF_tmt6        CGGGACCTCG CTCAGCTTCG CCGCCAGCGT TCACGGCCGC TGGCTGGTCG 349
Am_PC_tmt6        CGGGACCTCG CTCAGCTTCG CCGCCAGCGT TCACGGCCGC TGGCTGGTCG 349
Am_SH_tmt6        CGGGACCTCG CTCAGCTTCG CCGCCAGCGT TCACGGCCGC TGGCTGGTCG 349
Clustal Consensus *** ***:** ********** * ***** .*  *.******  ******** *

                                                        Exon 1 / Exon 2
Dr_tmt6           GGAGACACGG CTGCATGTGG TACGGCTTCG TCAACTCATG CTTTGGGATT 372
Am_SF_tmt6        GCCGGAGGGG CTGCATGTGG TACGGATTCG TCAACTCCTG CTTCGGAATC 399
Am_PC_tmt6        GCCGGAGGGG CTGCATGTGG TACGGATTCG TCAACTCCTG CTTCGGAATC 399
Am_SH_tmt6        GCCGGAGGGG CTGCATGTGG TACGGATTCG TCAACTCCTG CTTCGGAATC 399
Clustal Consensus * .*... ** ********** *****.**** *******.** *** **.**
```

```
Dr_tmt6           GTATCGCTGA TTTCTCTGGC AATCCTTTCG TATGATCGTT ACAGTACATT 422
Am_SF_tmt6        GTATCTCTGA TCTCTCTGGT GGTTCTATCG TATGAGAGAT ACAGCACTTT 449
Am_PC_tmt6        GTATCTCTGA TCTCTCTGGT GGTTCTATCG TATGAGAGAT ACAGCACTTT 449
Am_SH_tmt6        GTATCTCTGA TCTCTCTGGT GGTTCTATCG TATGAGAGAT ACAGCACTTT 449
Clustal Consensus ***** **** * *******  ..* **:*** ***** .*:* **** **:**

Dr_tmt6           AACTGTTTAC AACAAGAGGG CTCCAGACTA CAGTAAACCC TTGTTGGCAG 472
Am_SF_tmt6        AACAGTTTAT AATAAACAGG CCCCGGACTA CAGGAAGCCC CTGCTGGCGG 499
Am_PC_tmt6        AACGGTTTAT AATAAACAGG CCCCGGACTA CAGGAAGCCC CTGCTGGCGG 499
Am_SH_tmt6        AACGGTTTAT AATAAACAGG CCCCGGACTA CAGGAAGCCC CTGCTGGCGG 499
Clustal Consensus *** *****  ** **...** * **.***** *** **.***  ** ****.*

Dr_tmt6           TCGGGGGCTC GTGGCTCTAC TCACTGTTCT GGACGGTTCC TCCTCTGCTG 522
Am_SF_tmt6        TCGGGGGCTC CTGGCTGTAT TCGCTGCTGT GGACGGTGCC CCCCCTGCTG 549
Am_PC_tmt6        TCGGGGGCTC CTGGCTGTAT TCGCTGCTGT GGACGGTGCC CCCCCTGCTG 549
Am_SH_tmt6        TCGGGGGCTC CTGGCTGTAT TCGCTGCTGT GGACGGTGCC CCCCCTGCTG 549
Clustal Consensus **********  ***** **  **.*** * * ******* **  ** ******

Dr_tmt6           GGCTGGAGCA GCTATGGACT GGAAGGGGCA GGAACCAGCT GTTCGGTGAC 572
Am_SF_tmt6        GGCTGGAGTA GTTATGGGCT GGAAGGGGCA GGAACCAGCT GTTCCGTTAC 599
Am_PC_tmt6        GGCTGGAGTA GTTATGGGCT GGAAGGGGCA GGAACCAGCT GTTCCGTTAC 599
Am_SH_tmt6        GGCTGGAGTA GTTATGGGCT GGAAGGGGCA GGAACCAGCT GTTCCGTTAC 599
Clustal Consensus ******** * * *****.** ********** ********** **** ** **

Dr_tmt6           ATGGACTGCC AACACGCCCC AGTCGCATTC CTACATCATC TGCCTGTTCA 622
Am_SF_tmt6        CTGGACGTCC AAAACCGTGC AGTCCCATTC CTACATCATC TGCCTCTTCA 649
Am_PC_tmt6        CTGGACGTCC AAAACCCTGC AGTCCCATTC CTACATCATC TGCCTCTTCA 649
Am_SH_tmt6        CTGGACGTCC AAAACCCTGC AGTCCCATTC CTACATCATC TGCCTCTTCA 649
Clustal Consensus .*****  ** **.**    * **** ***** ********** ***** ****

Dr_tmt6           TTTTTTGCTT AGGGATCCCT GTGCTGGTCA TGGTGTACTG TTACAGTCGG 672
Am_SF_tmt6        TCTTCTGCCT CGGGATCCCA GTACTGATCA TGATGTACTG CTACAGCCGG 699
Am_PC_tmt6        TCTTCTGCCT CGGGATCCCA GTACTGATCA TGATGTACTG CTACAGCCGG 699
Am_SH_tmt6        TCTTCTGCCT CGGGATCCCA GTACTGATCA TGATGTACTG CTACAGCCGG 699
Clustal Consensus * ** *** * .********: **.***.*** **.*******  ***** ***

                                  Exon 2 / Exon 3
Dr_tmt6           CTCATCTGTG CTGTCAAACA GGTAGGCCGC ATTAGGAAGA CAGCAGCAAG 722
Am_SF_tmt6        CTGCTCTGCG CCGTAAAACA GGTTGGGCGC TTCAGGAAGA CGGCGGCGCG 749
Am_PC_tmt6        CTGCTCTGCA CCGTAAAACA GGTTGGGCGC TTCAGGAAGA CGGCAGCGCG 749
Am_SH_tmt6        CTGCTCTGCG CCGTAAAACA GGTTGGGCGC TTCAGGAAGA CGGCGGCGCG 749
Clustal Consensus ** .**** . * **.***** ***:** *** :* ******* *.**.**..*

Dr_tmt6           ACGGAGAGAG TACCACATTT TATTTATGGT CATCACTACA GTGGTGTGCT 772
Am_SF_tmt6        GAGGCGGGAG TACCACATCC TCTTCATGGT CGTCACCACG GTGGTGTGCT 799
Am_PC_tmt6        GAGGCGGGAG TACCACATCC TCTTCATGGT CGTCACCACG GTGGTGTGCT 799
Am_SH_tmt6        GAGGCGGGAG TACCACATCC TCTTCATGGT CGTCACCACG GTGGTGTGCT 799
Clustal Consensus ..**.*.*** ********   *.** ***** *.**** **. **********

Dr_tmt6           ACCTTTTGTG CTGGATGCCT TATGGAGTTG TTGCAATGAT GGCCACGTTC 822
Am_SF_tmt6        ATCTGGTGTG CTGGATGCCG TACGGTGTGG TTGCCATGAC GGCAACCTTC 849
Am_PC_tmt6        ATCTGGTGTG CTGGATGCCG TACGGTGTGG TTGCCATGAC GGCAACCTTC 849
Am_SH_tmt6        ATCTGGTGTG CTGGATGCCG TACGGTGTGG TTGCCATGAC GGCAACCTTC 849
Clustal Consensus * **  **** *********  ** **:** * ****.****  ***.** ***
```

```
Dr_tmt6           GGACGCCCAG GAATCATCTC ACCCATTGCA AGCGTGGTGC CATCCCTCCT 872
Am_SF_tmt6        GGCCAACCGG GGCTGATCAC CCCGGTGGTT AGCGTGGTGC CGTCACTACT 899
Am_PC_tmt6        GGCCGACCGG GGCTGATCAC CCCGGTGGTT AGCGTGGTGC CGTCACTACT 899
Am_SH_tmt6        GGCCGACCGG GGCTGATCAC CCCGGTGGTT AGCGTGGTGC CGTCACTACT 899
Clustal Consensus **.*..**.* *..* ***:* .** .* * : ********** *.**.**.**


Dr_tmt6           TGCCAAAAGC AGCACAGTCA TCAATCCTCT CATTTACATC CTCATGAACA 922
Am_SF_tmt6        GGCCAAGAGC AGCACCGTCT TCAACCCCAT CATTTACATC CTCATGAACA 949
Am_PC_tmt6        GGCCAAGAGC AGCACCGTCT TCAACCCCAT CATTTACATC CTCATGAACA 949
Am_SH_tmt6        GGCCAAGAGC AGCACCGTCT TCAACCCCAT CATTTACATC CTCATGAACA 949
Clustal Consensus  *****.*** *****.***: **** ** .* ********** **********

                  Exon 3 / Exon 4
Dr_tmt6           AACAGTTTTA CAGGTGTTTC CTCATCTTGA TTCACTGCAA ACATAGCTCT 972
Am_SF_tmt6        AACAGTTCTA CAGGTGTTTC CTGATCCTGT TTCACTGTAA ACACAGTTCA 999
Am_PC_tmt6        AACAGTTCTA CAGGTGTTTC CTGATCCTGT TTCACTGTAA ACACAGTTCA 999
Am_SH_tmt6        AACAGTTCTA CAGGTGTTTC CTGATCCTGT TTCACTGTAA ACACAGTTCA 999
Clustal Consensus ******* ** ********** ** *** **: ******* ** *** ** **:


Dr_tmt6           CTAGAGAACG GACAGTCATC AATGCCTTCA AGAACTACAG GCATCCAGCT 1022
Am_SF_tmt6        CACCTGAACG GTCACTC--- CATGCCCTCC AGAACTACGG TCATCCAGCT 1046
Am_PC_tmt6        CACCTGAACG GTCACTC--- CATGCCCTCC AGAACTACGG TCATCCAGCT 1046
Am_SH_tmt6        CACCTGAACG GTCACTC--- CATGCCCTCC AGAACTACGG TCATCCAGCT 1046
Clustal Consensus *:. :***** *:** **    .***** **. ********.*  *********


Dr_tmt6           TAATCGAAGG CCTTACAGCA ACCCAGTGGC TGACAACGCA CCCCCTTCCA 1072
Am_SF_tmt6        GAACCGGCGG CTCTGCAGCA ACACGGTAAC CGGAAACATC CCCACCTCCC 1096
Am_PC_tmt6        GAACCGGCGG CTCTGCAGCA ACACGGTAAC CGGAAACATC CCCACCTCCC 1096
Am_SH_tmt6        GAACCGGCGG CTCTGCAGCA ACACGGTAAC CGGAAACATC CCCACCTCCC 1096
Clustal Consensus  ** **..** *  *.***** **.*.**..*  *..***. . ***.* ***.


Dr_tmt6           TCGATCTCCA AAACGATTGC AGCACCCCTG TCTCTGGCTA A--------- 1113
Am_SF_tmt6        TCAGACTGAA CACCGAGATC AGCACCCCCG TCTCCGACCG GACCCACCCC 1146
Am_PC_tmt6        TCAGACTGAA CACCGAGATC AGCACCCCCG TCTCCGACCG GACCCACCCC 1146
Am_SH_tmt6        TCAGACTGAA CACCGAGATC AGCACCCCCG TCTCCGACCG GACCCACCCC 1146
Clustal Consensus **..:** .* .*.*** : * ******** * **** *.* . .


Dr_tmt6           ---------- --------
Am_SF_tmt6        CCCCAGATCA CTCCTTGA 1164
Am_PC_tmt6        CCCCAGATCA CTCCCTGA 1164
Am_SH_tmt6        CCTCAGATCA CTCCTTGA 1164
Clustal Consensus
```

(B) Partial *tmt9*

```
Dr_tmt9           ATGTTTTTCG AGCAGGCCGA TTTAAACTAC AGCTTCAACA TGAGCGAGGA 50
Am_PC_tmt9L       ---------- ---------- ---------- ---------- ----------
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt9           AGACCGCTTA ACTCTTCTGG ACGAGGACTG GAGCGACTCG CCAATGGAGA 100
Am_PC_tmt9L       ---------- ---------- ---------- ---------- ----------
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt9           CGCTCTCACG CGCGGGCTTC ATCGCGCTCT CTGTGTTCCT GGGATTCATC 150
Am_PC_tmt9L       ---------- ---------- ---------- ---------- ----------
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt9           ATGACTTTCG GCTTTTTTAA TAACCTCGTA GTATTAGTGC TTTTCTGTAA 200
Am_PC_tmt9L       ---------- ---------- ---------- ---------- ----------
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt9           GTTTAAGACG CTGAGGACGC CGGTTAACAT GCTGCTTCTT AACATCAGCA 250
Am_PC_tmt9L       ---------- ---------- ---------- ---------- --------CG 2
Am_PC_tmt9S       ---------- ---------- ---------- ---------- --------CG 2
Clustal Consensus                                                      *.

Dr_tmt9           TCAGTGACAT GTTGGTGTGT ATGTTCGGGA CGACCCTGAG CTTCGCGTCC 300
Am_PC_tmt9L       TCAGTGACAT GCTGGTCTGC ACGTGCGGGA CGACGCTCAG CTTCGCGTCC 52
Am_PC_tmt9S       TCAGTGACAT GCTGGTCTGC ACGTGCGGGA CGACGCTCAG CTTCGCGTCC 52
Clustal Consensus ********** * **** **  * ** ***** **** ** ** **********

Dr_tmt9           AGTGTGCGGG GAAGGTGGCT GCTCGGGCGA CACGGCTGCA TGTGGTACGG 350
Am_PC_tmt9L       AGCATCCACG GCCGGTGGCT GCTTGGGAGG CAGGGCTGCA TGTGGTACGG 102
Am_PC_tmt9S       AGCATCCACG GCCGGTGGCT GCTTGGGAGG CAGGGCTGCA TGTGGTACGG 102
Clustal Consensus ** .* *. * *..******* *** ***.*. ** ******* **********

                                  Exon 1 / Exon 2
Dr_tmt9           CTTCATCAAC TCCTGCTTCG|GTATTGTTTC TCTCATATCC TTGGTGGTTC 400
Am_PC_tmt9L       ATTCATCAAC TCCTGCTTCG|GTATCGTATC TCTAATATCC TTGGTAATTC 152
Am_PC_tmt9S       ATTCATCAAC TCCTGCTTCG|GTATCGTATC TCTAATATCC TTGGTAATTC 152
Clustal Consensus .********* **********|**** **:** ***.****** *****..***

Dr_tmt9           TTTCTTACGA CCGCTACAGC ACATTGACTG TGTACCATAA ACGGGCTCCA 450
Am_PC_tmt9L       TGTCGTACGA CCGCTACAGC ACGCTGACGG TGTACAACAA GAAGGGTCCA 202
Am_PC_tmt9S       TGTCGTACGA CCGCTACAGC ACGCTGACGG TGTACAACAA GAAGGGTCCA 202
Clustal Consensus * ** ***** ********** **. **** * *****.* ** ...** ****

Dr_tmt9           GACTACCGCA AGCCTCTGCT GGCAGTAGGA GGCTCCTGGC TGTACTCTCT 500
Am_PC_tmt9L       GACTACCGTA AGCCCCTGCT GGCAGTGGGG GGCTCGTGGC TGTACTCTGT 252
Am_PC_tmt9S       GACTACCGTA AGCCCCTGCT GGCAGTGGGG GGCTCGTGGC TGTACTCTGT 252
Clustal Consensus ******** * **** ***** ******.**. ***** **** ******** *
```

```
Dr_tmt9            GATCTGGACA GTGCCTCCTC TTTTAGGCTG GAGCAGTTAT GGGCTGGAGG 550
Am_PC_tmt9L        GGTCTGGACC GTGCCTCCTC TGCTGGGCTG GAGTAGCTAT GGTCTAGAAG 302
Am_PC_tmt9S        GGTCTGGACC GTGCCTCCTC TGCTGGGCTG GAGTAGCTAT GGTCTAGAAG 302
Clustal Consensus  *.*******. ********** *  *.***** *** ** *** ** **.**.*

Dr_tmt9            GTGCCGGAAC CAGCTGCTCT GTATCATGGA CTCAGCGCAC GGCCGAGTCT 600
Am_PC_tmt9L        GGGCAGGAAC CAGCTGTTCT GTATCATGGA CCGAGCACTC ACCTCAGTCT 352
Am_PC_tmt9S        GGGCAGGAAC CAGCTGTTCT GTATCATGGA CCGAGCACTC ACCTCAGTCT 352
Clustal Consensus  * **.***** ****** *** ********** *  ***.*:* . *  *****

Dr_tmt9            CATGCCTACA TCATCTGTCT GTTTGTATTC TGCCTGGGAC TACCAGTGCT 650
Am_PC_tmt9L        CATGCCTACA TCATCTGTCT GTTCATCTTC TGCCTGGCTC TGCCTGTGCT 402
Am_PC_tmt9S        CATGCCTACA TCATCTGTCT GTTCATCTTC TGCCTGGCTC TGCCTGTGCT 402
Clustal Consensus  ********** ********** *** .*.*** ******* :* *.**:*****

                                                          Exon 2 / Exon 3
Dr_tmt9            TGTCATGGTC TACTGCTATG GGAGACTGCT GTACGCCGTC AAACAGGTAG 700
Am_PC_tmt9L        GGTCATGGTC TACAGCTATG GGAGGCTACT GTACGCAGTC AAACAGGTAG 452
Am_PC_tmt9S        GGTCATGGTC TACAGCTATG GGAGGCTACT GTACGCAGTC AAACAGGTAG 452
Clustal Consensus   ********* ***:****** ****.**.** ******.*** **********

Dr_tmt9            GGAAGATACG CAAGACGGCT GCACGGAAAC GGGAGTACCA TGTGTTGTTC 750
Am_PC_tmt9L        GAAAGATTCG CAAGTCTGCA GCTCGGAGGA GGGAGTATCA CGTGTTGTTT 502
Am_PC_tmt9S        GAAAGATTCG CAAGTCTGCA GCTCGGAGGA GGGAGTATCA CGTGTTGTTT 502
Clustal Consensus  *.*****:** ****:* **: **:****... ******* **  ********

Dr_tmt9            ATGGTTATCA CCACTGTGGT GTGTTACCTT TTGTGCTGGA TGCCGTACGG 800
Am_PC_tmt9L        ATGGTGATCA CTACAGTGGT GTGTTACCTG CTGTGCTGGA CGCCGTACAG 552
Am_PC_tmt9S        ATGGTGATCA CTACAGTGGT GTGTTACCTG CTGTGCTGGA CGCCGTACAG 552
Clustal Consensus  ***** **** * **:***** *********   *********  *******.*

Dr_tmt9            CGTCGTGGCC ATGATGGCCA CATTTGGCCG GCCGGGAATC ATCTCACCTG 850
Am_PC_tmt9L        CGTGGTGGCC ATAATGGCCA CGTTTGGTCG GCCAGGAATC ATCACACCTG 602
Am_PC_tmt9S        CGTGGTGGCC ATAATGGCCA CGTTTGGTCG GCCAGGAATC ATCACACCTG 602
Clustal Consensus  *** ****** **.******* *.***** ** ***.****** ***:******

Dr_tmt9            TGGCCAGTGT GGTTCCCTCT CTCCTCGCCA AGAGCAGCAC GGTCATCAAC 900
Am_PC_tmt9L        TGGCGAGCGT GGTGCCCTCT CTCCTAGCCA AGAGCAGCAC AGTCATCAAC 652
Am_PC_tmt9S        TGGCGAGCGT GGTGCCCTCT CTCCTAGCCA AGAGCAGCAC AGTCATCAAC 652
Clustal Consensus  **** ** ** *** ****** *****.**** ********** .*********

                                                  Exon 3 / Exon 4
Dr_tmt9            CCACTCATAT ACATCCTCAT GAACAAACAG TTCTACAGGT GTTTCCGCAT 950
Am_PC_tmt9L        CCAGTTATTT ACATTCTTAT GAACAAACAG TTCTACAGGT GTTTCCTGAT 702
Am_PC_tmt9S        CCAGTTATTT ACATTCTTAT GAACAAACAG GTAAGAGCTT TCATTCATTC 702
Clustal Consensus  *** * **:* **** ** ** ********** *.:...  *   :* *  :

Dr_tmt9            CTTGTTCTGC TGTCAGAGAA GTTTGCTACA GAACGGTCAC TCCTCCATGC 1000
Am_PC_tmt9L        CCTGTTCCAC TGCAGAGGGA GGTCTATGGA GAATGGTCAC TCGTCCATGC 752
Am_PC_tmt9S        ACCAAAGA-- --------CA AATCAATCTC TAATATTTAT TTATTTCTAA 742
Clustal Consensus  .  .::             * . *  .*  .  ** . * *  *  *  .*..
```

```
Dr_tmt9           CCTCCAAGAC CACCGTCATC CAGCTGAACC GCAGGGTCAA CAGTAACGCT 1050
Am_PC_tmt9L       AGTCTAAGAC GACGGTCATC CATCTGAACC GTGTGGTGCA CAACAACACT 802
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ----------
Clustal Consensus


Dr_tmt9           GTAGCCTGCA CCGCTCAGAT CTCCACCGGG ACTCACAACC ATGATTGCAG 1100
Am_PC_tmt9L       GTGGCCTGCC AGGCGCAGAT CTCCACCGGC CTGCAGGAAC AAGGCTTAAG 852
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ----------
Clustal Consensus


Dr_tmt9           CACTCACGTC ACAGAGAGGA GCAACCCTCC AGAGGTCATT CCATGA     1146
Am_PC_tmt9L       CACCACAGCC ATGGAGAGGA GCAACCCTCC AGAAATCTCG TCATAA     898
Am_PC_tmt9S       ---------- ---------- ---------- ---------- ------
Clustal Consensus
```

(C)Partial *tmt14*

```
Dr_tmt14          ATGGTCGTCT ACATCTGGAG TTTGAACATC AGCAGCAAAG ACACTTCTGC 50
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          GCTTAACCAA AGTGGCAATG TGAGTTCAGG GGATCCTCTG GAGCCCCATG 100
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          ACAGCCCACC CGGCCTGAGC AGGACCGGCC ACACGGTGAC CGCCGTCTGC 150
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          CTCGGAGCCA TCCTGCTGTT GGGATGCCTG AACAACCTCT TTGTCCTGCT 200
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          CGTCTTCGCG AGGTTTCGCA CACTGTGGAC TCCTATAAAC CTGATCCTGC 250
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          TGAACATCAG CGTCAGTGAC ATACTGGTCT GCTTGTTTGG GACTCCCTTT 300
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          AGCTTCGCGT CCAGTCTCTA TGGGAAATGG CTTTTGGGAC ATCACGGCTG 350
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus
                                                 Exon 1 / Exon 2
Dr_tmt14          CAAATGGTAC GGTTTCGCCA ATTCGCTTTT TGGAATTGTG TCTCTGATGT 400
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          CTCTTTCTAT ACTGTCATAC GAGCGTTATG CTGCCCTGCT GCGGGCCACC 450
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt14          AAGGCCGATG TGTCTGACTT CCGCAGGGCC TGGCTCTGTG TGGCTGGTTC 500
Am_PC_tmt14       ---------- ---------- ---------- ---------- ----------
Am_SH_tmt14       ---------- ---------- ---------- ---------- ----------
Clustal Consensus
```

```
Dr_tmt14          CTGGCTATAT TCACTGCTAT GGACTCTTCC ACCATTTCTG GGATGGAGCA 550
Am_PC_tmt14       ---------- ---------- ----TCTGCC TCCATTCTTA GGCTGGAGCA 26
Am_SH_tmt14       ---------- ---------- ----TCTGCC TCCATTCTTA GGCTGGAGCA 26
Clustal Consensus                          *** ** :*****  *. **.*******

Dr_tmt14          ACTATGGTCC TGAAGGACCC GGGACTACAT GCTCAGTGCA ATGGCACCTG 600
Am_PC_tmt14       GCTATGGGCC TGAAGGTCCA GGAACCACCT GTTCTGTGCA GTGGCACCAG 76
Am_SH_tmt14       GCTATGGGCC TGAAGGTCCA GGAACCACCT GTTCTGTGCA GTGGCACCAG 76
Clustal Consensus .****** ** ******:**. **.** **.* * **:***** .*******:*

Dr_tmt14          CGCTCAACCA GCAGCATCTC GTACGTCATG TGCTTGTTCA TCTTCTGTCT 650
Am_PC_tmt14       AGATCCCCCA GCAGCATCTC GTATGTGGTG TGCCTGTTCA TCTTCTGCCT 126
Am_SH_tmt14       AGATCCCCCA GCAGCGTCTC GTATGTGGTG TGCCTGTTCA TCTTCTGCCT 126
Clustal Consensus .*.**..*** *****.**** *** ** .** *** ****** ******* **

Dr_tmt14          GCTTCTGCCC CTCGTGCTCA TGATCTTCTG CTATGGCAAA ATCCTGCTAC 700
Am_PC_tmt14       CGTCCTGCCA CTGCTCCTCA TGGTGTACTG CTATGGCAAG ATCCTTTTTA 176
Am_SH_tmt14       CGTCCTGCCA CTGCTCCTCA TGGTGTACTG CTATGGCAAG ATCCTTTTTA 176
Clustal Consensus   * *****. **  * **** **.* *:*** *********. *****  *:.

                            Exon 2 / Exon 3
Dr_tmt14          TCATCAAAGG GGTAACAAAA ATCAACCTGC TGACTGCGCA GAGGAGAGAG 750
Am_PC_tmt14       TCATTAAAGG GGTGACTAAG ATCAATCTGC TGTCAGCACA AAGGCGGGAG 226
Am_SH_tmt14       TCATTAAAGG GGTGACTAAG ATCAATCTGC TGTCAGCACA AAGGCGGGAG 226
Clustal Consensus **** ***** ***.**:**. ***** **** **:*:**.** .***.*.***

Dr_tmt14          AACCACATTC TTCTGATGGT TGTCACTATG GTATCCTGCT ATCTGCTGTG 800
Am_PC_tmt14       AACCACATTC TACTGATGGT GCTGACGATG GTGTCCTGCT ACCTGCTGTG 276
Am_SH_tmt14       AACCACATTC TACTGATGGT GCTGACGATG GTGTCCTGCT ACCTGCTGTG 276
Clustal Consensus ********** *:********   * ** *** **.******* * ********

Dr_tmt14          CTGGATGCCT TACGGGGTGG TGGCCCTGCT GGCCACCTTT GGCAGAACGG 850
Am_PC_tmt14       CTGGATGCCC TATGGGGTGG TTTCGCTGAT GGCCACCTTT GGAAAGCAAG 326
Am_SH_tmt14       CTGGATGCCC TATGGGGTGG TTTCGCTGAT GGCCACCTTT GGAAAGCAAG 326
Clustal Consensus *********  ** ******* *  * ***.* ********** **.*.....*

Dr_tmt14          GACTAATCAC TCCTGTAACC AGCATAGTGC CATCAGTCCT GGCCAAGAGC 900
Am_PC_tmt14       GACTAATCAC GCCAGTTGCC AGCGTGGTGC CCTCAGTGTT AGCCAAGAGC 376
Am_SH_tmt14       GACTAATCAC GCCAGTTGCC AGCGTGGTGC CCTCAGTGTT AGCCAAGAGC 376
Clustal Consensus **********  **:**:.** ***.*.**** *.*****  * .*********

                                                      Exon 3 / Exon 4
Dr_tmt14          AGCACTGTGG TCAACCCAGT AATCTATGTG CTTTTCAACA ACCAGTTCTA 950
Am_PC_tmt14       AGCACAGTGA TCAATCCGGT TATATACGTG CTCTTCAACA ACCAGTTCTA 426
Am_SH_tmt14       AGCACAGTGA TCAATCCGGT TATATACGTG CTCTTCAACA ACCAGTTCTA 426
Clustal Consensus *****:***. **** **.** :**.** *** ** ******* **********

Dr_tmt14          CAGGTGTTTC GTAGCATTTC TGAAATGTCA AGGCGAACCA TCTGTCCATG 1000
Am_PC_tmt14       CAGGTGCTTC ATAGCTTTTG TAAAGTGTGG AGCAGAGCCC TCTATACAAA 476
Am_SH_tmt14       CAGGTGCTTC ATAGCTTTTG TAAAGTGTGG AGCAGAGCCC TCTATACAAA 476
Clustal Consensus ****** *** .****:***  *.**.*** . ** .**.**. ***.*.**:.
```

```
Dr_tmt14           GACAGAATCC ACAGCATAGC AGCAAAGAAG ATCCACACGT GTTCAGGCCC 1050
Am_PC_tmt14        CCCTGAACCC CCCACACAAC AGCAAGGAGG AAGCCCCTGC GTCCAGCTTT 526
Am_SH_tmt14        CCCTGAACCC CCCACACAAC AGCAAGGAGG AAGCCCCCGC GTCCAGCTTT 526
Clustal Consensus  .*:*** ** .*..** *.* *****.**.* *: *.*. *  ** ***


Dr_tmt14           TGTGACGGGG CCTCAATTCA TCGCAGCGCA GAGGGCCCTC AGAAAAAAGA 1100
Am_PC_tmt14        TGTAGACCCT GGAGCCACAG TAAGCTTTCC CCGGCGAGAG AG-----TGC 576
Am_SH_tmt14        TGTAGACCCT GGAGCCACAG TAAGCTTTCC CCGGCGAGAG AG-----TGC 576
Clustal Consensus ***...       : ..: .. *.. .   *.  .**  . :  **     :*.


Dr_tmt14           GCAGCACTCT CTGTCCTTGG TGGTTCAC-T ACACACCCTG A          1140
Am_PC_tmt14        CGCACCCTGT CCCTAGTGGT TCACTACACG CCCTGA---- -          607
Am_SH_tmt14        CGCACCCTGT CCCTAGTGGT TCACTACACG CCCTGA---- -          607
Clustal Consensus    ..*.** * *  *. * *  * . *...   .*. ..
```

(D) Partial *tmt24*

```
Dr_tmt24            ATGATTGAGT CTAACGTGAG TCGGAGTTGC GAGTGGTGCG CTGGAGGAGG 50
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            TGAAGGGACC GGCGCGCACT TGGATGAAAA TCACTCAGAC CACAGTCTGA 100
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            GCCCGACAGG ACACTTGGTA GTCGCGGTGT GCCTTGGATT CATCGGAACT 150
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            TTCGGCTTTC TCAACAACAC GCTCGTGCTC GTCCTGTTCT GTCGCTACAA 200
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            AGTGCTGCGC TCGCCTATGA ACTGTCTCCT GATCAGCATC TCAGTGAGCG 250
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            ATCTGCTGGT CTGCGTACTC GGCACTCCGT TCAGCTTCGC CGCCAGCACG 300
Am_PC_tmt24         ---------- ---------- ---------- ------TCGC CGCCAGCACG 14
Clustal Consensus                                         **** **********

Dr_tmt24            CAGGGTCGCT GGCTCATCGG GCGCGCGGGA TGCGTGTGGT ACGGTTTCAT 350
Am_PC_tmt24         CGCGGACGAT GGCTGATCGG CGGGGCTGGC TGCGTGTGGT ACGGCTTCGT 64
Clustal Consensus   *  ** ** * **** *****   * ** **  ********** **** *** *

                              Exon 1 / Exon 2

Dr_tmt24            CAATTCATTC CTGGGCGTTG TGTCTCTAAT CTCTCTGGCG GTTCTCTCAT 400
Am_PC_tmt24         CAACTCCTTC CTGGGTATCG TGTCTCTGAT CTCCCTGGCC GTGCTGTCCT 114
Clustal Consensus   *** ** *** *****  * * ******* ** *** *****  ** ** ** *

Dr_tmt24            ACGAGCGTTA CTGCACCATG ATGGGCTCCA CACAGGCCGA CTCCACCAAC 450
Am_PC_tmt24         ATGAGCGCTA CTGCACAATG ATGGGGGCCA CACAGGCCGA CTCCACCAAC 164
Clustal Consensus   * ***** ** ****** *** *****  *** ********** **********

Dr_tmt24            TACCGAAAAG TGGTCATAGG CATTGCTTTC TCCTGGATTT ATTCCATGGT 500
Am_PC_tmt24         TACAGAAAAG TGATCATGGG AATCACGTTC TCCTGGATTT ACTCCATGAT 214
Clustal Consensus   *** ****** ** **** **  **  * *** ********** * ****** *

Dr_tmt24            TTGGACTCTT CCACCACTGT TCGGATGGAG CTGCTACGGT CCGGAGGGTC 550
Am_PC_tmt24         CTGGACTCTG CCCCCGCTGT TCGGCTGGAG CCGCTATGGC CCCGAGGGGC 264
Clustal Consensus    ********  ** ** **** **** ***** * **** **  ** ***** *

Dr_tmt24            CCGGCACCAC CTGCAGCGTC AACTGGGCCG CGCGGACTCC CAACAACGTG 600
Am_PC_tmt24         CGGGCACTAC CTGCAGTGTG AACTGGGCTG CTAAGACAGC CAACAATGTG 314
Clustal Consensus   * ***** ** ****** **  ******** * *   ***  * ****** ***
```

```
Dr_tmt24            TCTTACATTG TCTGTCTGTT TGTCTTCTGC CTTATTCTCC CTTTTATCGT 650
Am_PC_tmt24         TCTTACATCA TCTGCCTCTT CTTCTTCTGC CTCATCCTTC CCTTTTTTGT 364
Clustal Consensus   ********   **** ** **   ******** ** ** ** * * *** * **
```

Exon 2 / Exon 3

```
Dr_tmt24            GATCGTCTAC AGCTACGGCA GACTCCTTCA GGCTATCACA CAGGTGAGCA 700
Am_PC_tmt24         GATCATCTAC AGCTATGGGA AGCTTCTACA AGCCATCAAG CAGGTGAGCA 414
Clustal Consensus   **** ***** ***** ** *   ** ** **  ** ****   ********** 

Dr_tmt24            GGATCAATAC GGTGGTGAGT CGTAAACGGG AGCAGCGTGT GCTCTTCATG 750
Am_PC_tmt24         GGATCAACAC TGTGGTGAGT CGTAAGCGGG AGCAGCGTGT GCTCTTATTG 464
Clustal Consensus   ******* **  ********* ***** **** ********** ******  ** 

Dr_tmt24            GTGGTCACCA TGGTGGTCTG TTACCTGCTT TGCTGGTTGC CATATGGGAT 800
Am_PC_tmt24         GTGATCACCA TGGTGGTGTG TTACCTGTTG TGCTGGCTGC CCTACGGCAT 514
Clustal Consensus   *** ****** ******* ** ******* *  ****** *** * ** ** ** 

Dr_tmt24            AATGGCGCTG CTGGCTACGT TTGGACATCC CGGACTGGTG ACACCAGCAG 850
Am_PC_tmt24         CATGGCCTTG GTGGCCACCT TCGGGCATCC TGGCCTGGTC ACTCCAGAAG 564
Clustal Consensus    *****  **  **** ** * * ** *****  ** *****  ** **** ** 

Dr_tmt24            CCAGCATAGT TCCTTCTCTG TTAGCCAAGA GCAGCACCGT CATAAACCCA 900
Am_PC_tmt24         CCAGCATTGT TCCCTCATTA CTCGCCAAAT CCAGCACAGT CATCAACCCA 614
Clustal Consensus   ******* ** *** **  *   * *****    ****** ** *** ****** 
```

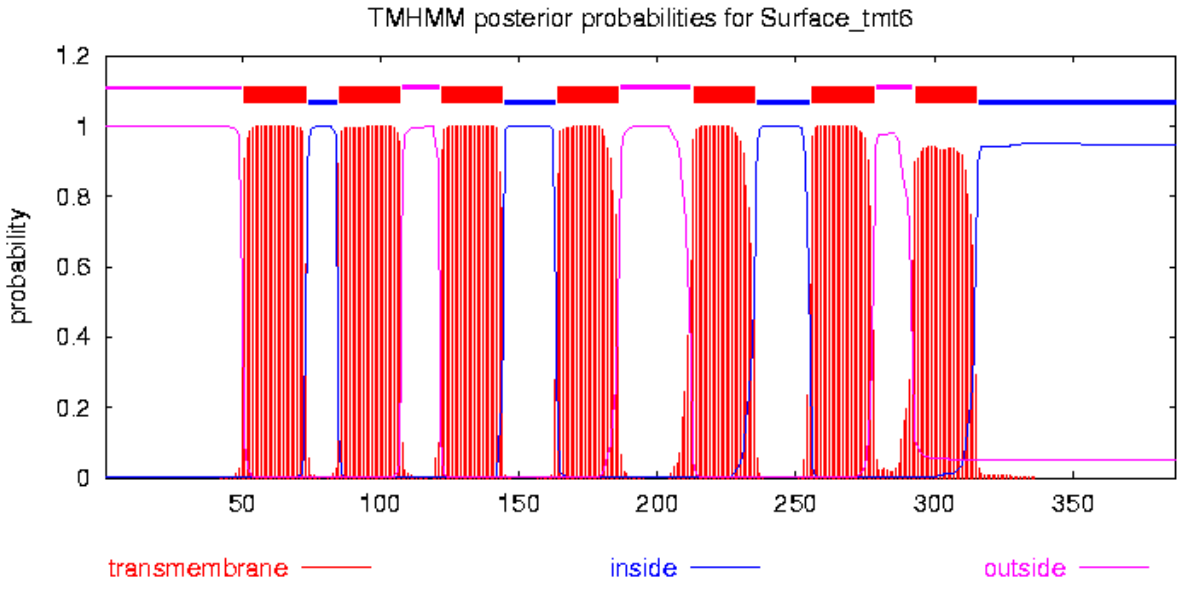
Exon 3 / Exon 4

```
Dr_tmt24            ATCATCTACA TCTTCATGAA CAAACAGTTT TGCAGATGTT TCCATGCGCT 950
Am_PC_tmt24         GTCATCTACA TCTTC----- ---------- ---------- ---------- 629
Clustal Consensus    ********* *****

Dr_tmt24            CATCATGTGC ACCACCCCAG AGAGAGGCTC CAGCTTTAAA AACTCTTCCA 1000
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            AAGTCACCAA AACTCTCAGG ACAGTGCGAC GTGCCAACGG CCAGAATGTG 1050
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            ACCTTCGCTG TGGCATCGGC CGTCCATCGA ACTCCGTACA GCGACAGACA 1100
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            GAAAAGTTCC TCGGAGGGTG AGAAACTCCC TCCAGCTACG GGTCAGGGCA 1150
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----------
Clustal Consensus

Dr_tmt24            CCAGCAAACC TGTTGTCTCT CTTGTGGCCT ATTACAATGG CTGA       1200
Am_PC_tmt24         ---------- ---------- ---------- ---------- ----
Clustal Consensus
```

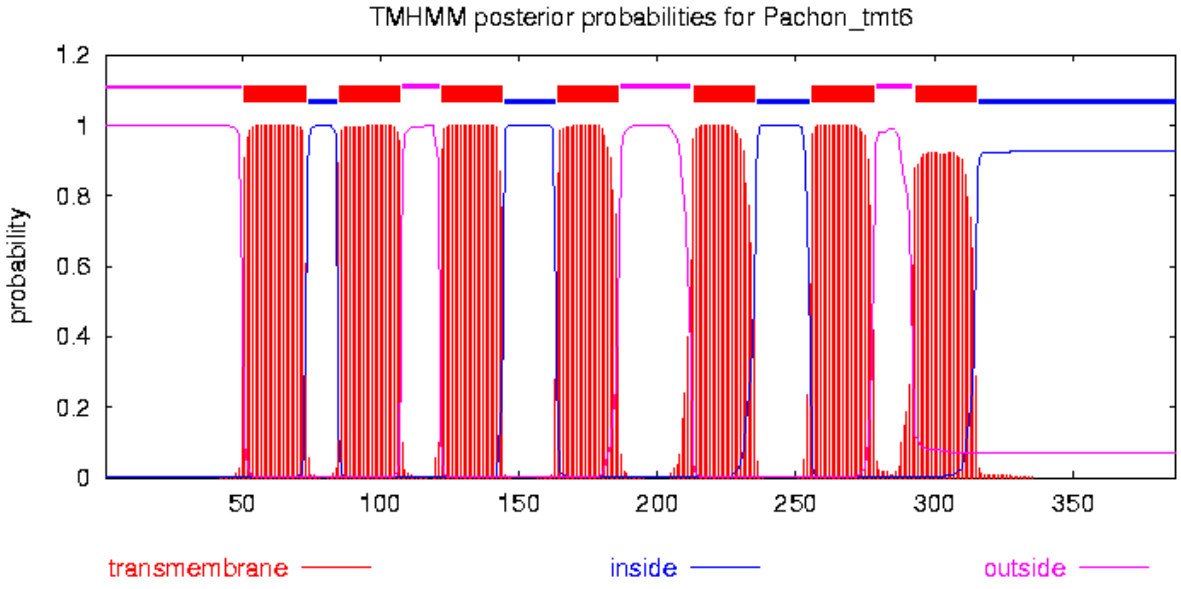
Figure A.3 Alignment of *Astyanax mexicanus* (*Am*) *tmt* opsin nucleotide sequences with zebrafish (*Danio rerio, Dm*) *tmt* opsin sequences. Exon-exon boundaries are predicted based on the zebrafish *tmt* opsin sequences. Gaps are inserted as dashes (-) to maintain a high level of identity between the sequences. Symbols below each set of the alignment delineates level of conservation between the sequences: * = consensus residue between all the sequences examined; := different aa but of similar charges across the sequences; .= different aa with distinctive charges across the sequences.

A.4 Prediction of transmembrane domains for *Astyanax mexicanus* tmt opsins

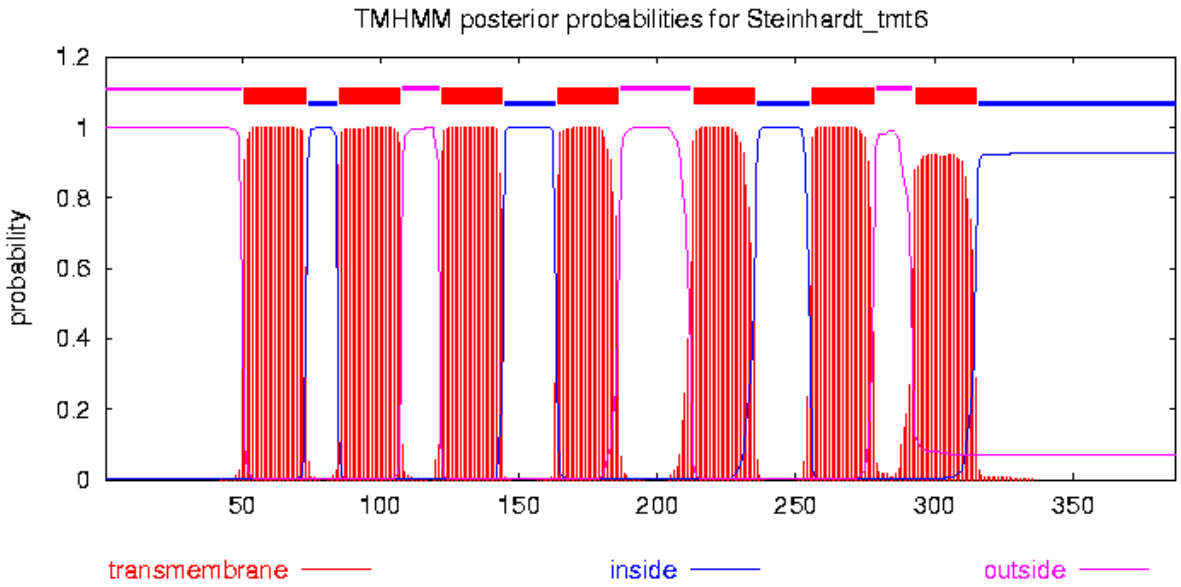
(A) Surface tmt6



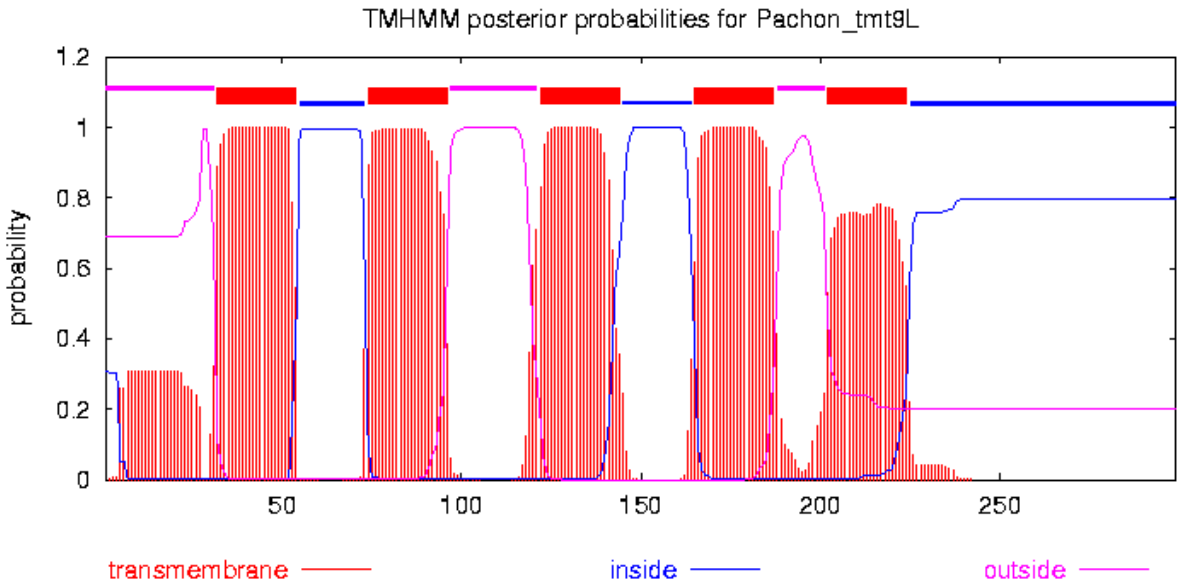
(B) Pachón tmt6



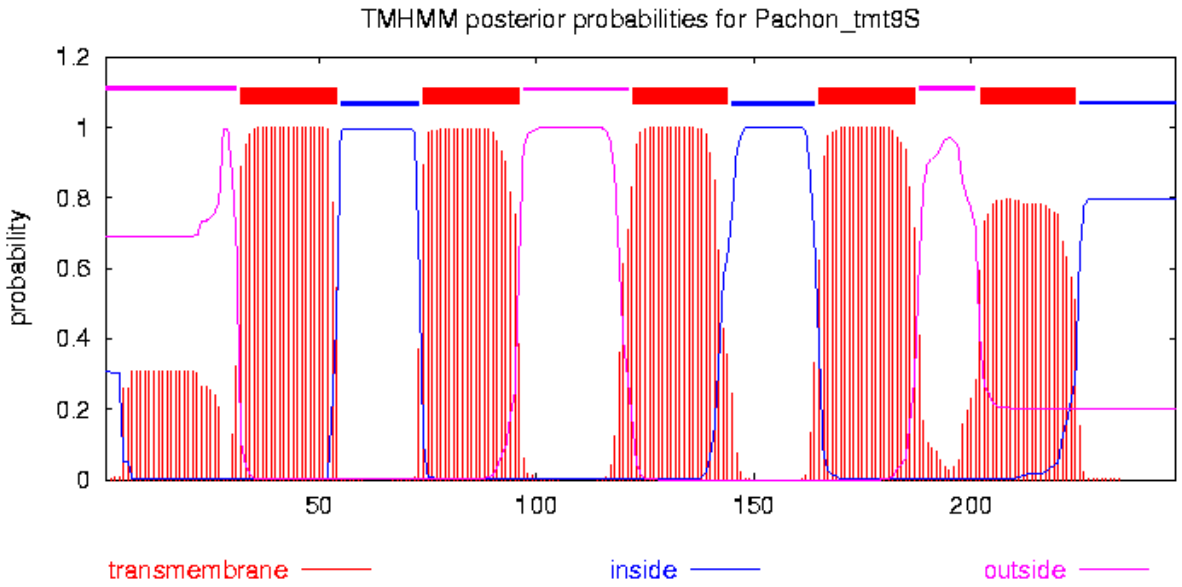
(C) Steinhardt tmt6



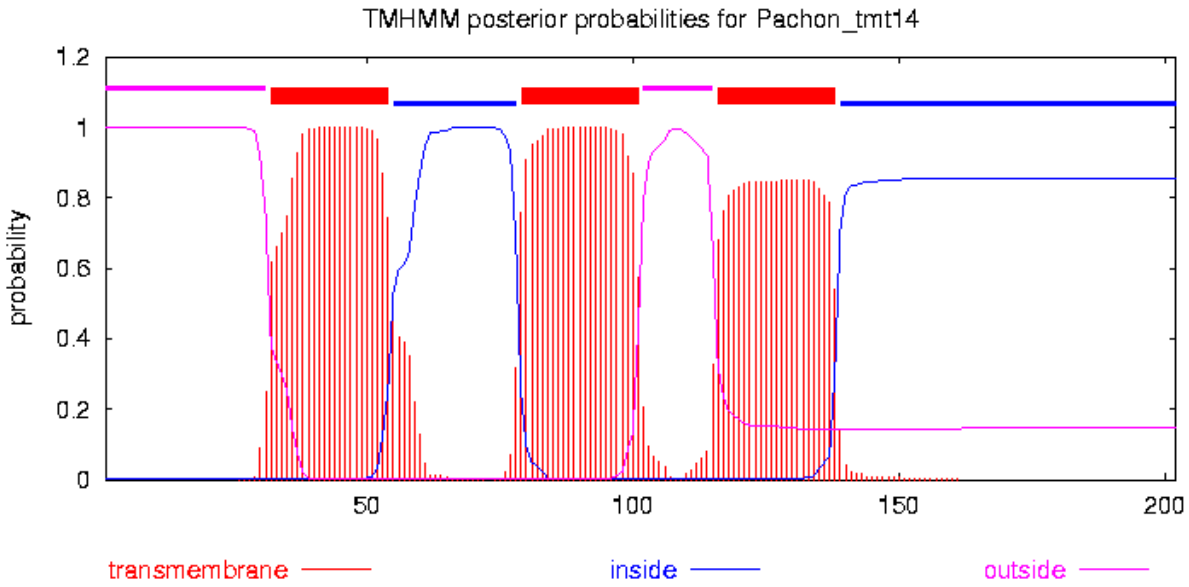
(D) Pachón tmt9l (partial)



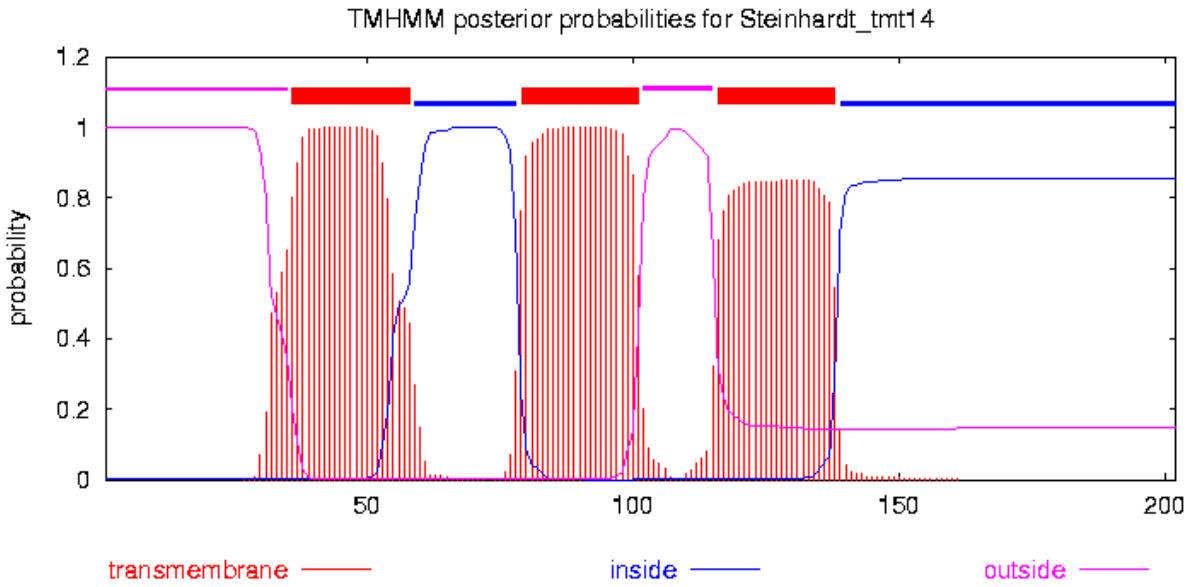
(E) Pachón tmt9s (partial)



(F) Pachón tmt14 (partial)



(G) Steinhardt tmt14 (partial)



(H) Pachón tmt24 (partial)

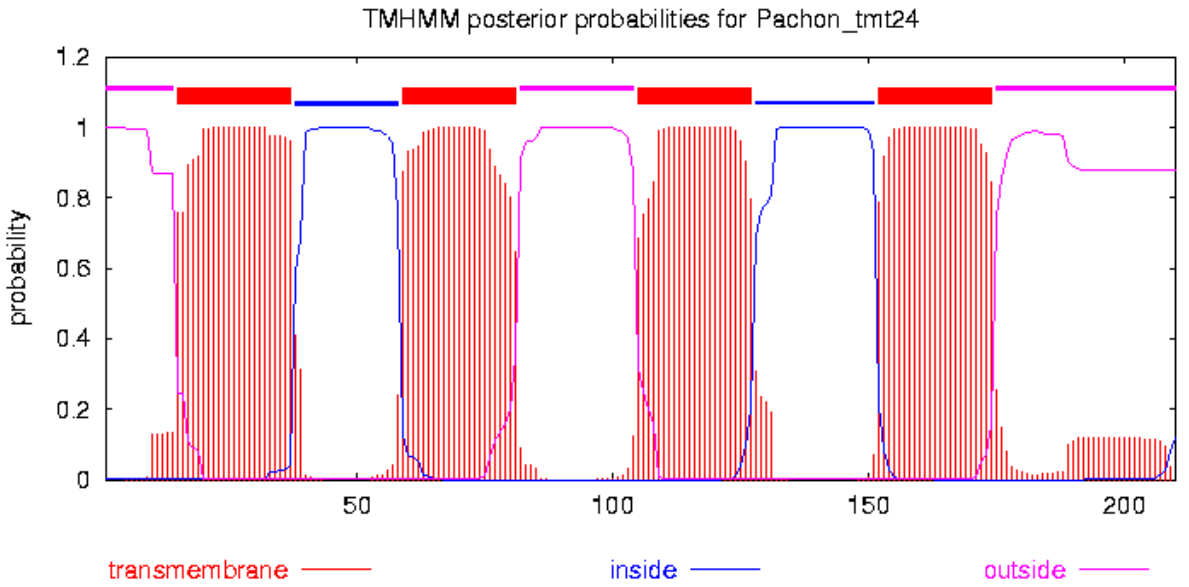


Figure A.4 Hydrophobicity plots of *Astyanax mexicanus* tmt opsins. The deduced amino acid sequences of **(A)** Surface tmt6, **(B)** Pachón tmt6, **(C)** Steinhardt tmt6, **(D)** Pachón tmt9l (partial), **(E)** Pachón tmt9s (partial), **(F)** Pachón tmt14 (partial), **(G)** Steinhardt tmt14 (partial), and **(H)** Pachón tmt24 (partial) were analysed using an online program TMHMM Server v2.0 (http://www.cbs.dtu.dk/services/TMHMM/) for prediction of hydrophobic transmembrane regions. Probability of the tmt sequences forming transmembrane domains are indicated in red lines, ranging between 0 - 1.0 on the y-axis. The intracellular domains are shown in blue lines, whilst the extracellular domains are shown in pink lines. Seven transmembrane helices were predicted for full-length sequences of Surface, Pachón and Steinhardt tmt6; five transmembrane helices were identified for the partial sequences of Pachón tmt9l and tmt9s; four transmembrane helices were identified for the partial sequence of Pachón tmt24; and three transmembrane helices were observed for the partial sequences of Pachón and Steinhardt tmt14.

References

2002. Ultraviolet radiation related exposures: broad-spectrum ultraviolet (UV) radiation, UVA, UVB, UVC, solar radiation, and exposure to sunlamps and sunbeds. *Rep Carcinog,* 10**,** 250-4.

AKIMOTO, M. 2005. Transcriptional factors involved in photoreceptor differentiation. *Semin Ophthalmol,* 20**,** 25-30.

AL-JANDAL, N., FARRAR, G. J., KIANG, A. S., HUMPHRIES, M. M., BANNON, N., FINDLAY, J. B., HUMPHRIES, P. & KENNA, P. F. 1999. A novel mutation within the rhodopsin gene (Thr-94-Ile) causing autosomal dominant congenital stationary night blindness. *Hum Mutat,* 13**,** 75-81.

ALA-LAURILA, P., ALBERT, R. J., SAARINEN, P., KOSKELAINEN, A. & DONNER, K. 2003. The thermal contribution to photoactivation in A2 visual pigments studied by temperature effects on spectral properties. *Vis Neurosci,* 20**,** 411-9.

ALA-LAURILA, P., KOLESNIKOV, A. V., CROUCH, R. K., TSINA, E., SHUKOLYUKOV, S. A., GOVARDOVSKII, V. I., KOUTALOS, Y., WIGGERT, B., ESTEVEZ, M. E. & CORNWALL, M. C. 2006. Visual cycle: Dependence of retinol production and removal on photoproduct decay and cell morphology. *Journal of General Physiology,* 128**,** 153-169.

ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2002. Signaling through G-Protein-Linked Cell-Surface Receptors. *Molecular Biology of the Cell.* 4th Edition ed. New York: Garland Science.

ALI, H. & MURRELL, J. C. 2009. Development and validation of promoter-probe vectors for the study of methane monooxygenase gene expression in Methylococcus capsulatus Bath. *Microbiology,* 155**,** 761-71.

ALLEN, A. E., BROWN, T. M. & LUCAS, R. J. 2011. A distinct contribution of short-wavelength-sensitive cones to light-evoked activity in the mouse pretectal olivary nucleus. *J Neurosci,* 31**,** 16833-43.

ALLISON, W. T., BARTHEL, L. K., SKEBO, K. M., TAKECHI, M., KAWAMURA, S. & RAYMOND, P. A. 2010. Ontogeny of cone photoreceptor mosaics in zebrafish. *J Comp Neurol,* 518**,** 4182-95.

ALTIMUS, C. M., GULER, A. D., ALAM, N. M., ARMAN, A. C., PRUSKY, G. T., SAMPATH, A. P. & HATTAR, S. 2010. Rod photoreceptors drive circadian photoentrainment across a wide range of light intensities. *Nat Neurosci,* 13**,** 1107-12.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

ALUNNI, A., MENUET, A., CANDAL, E., PENIGAULT, J. B., JEFFERY, W. R. & RETAUX, S. 2007. Developmental mechanisms for retinal degeneration in the blind cavefish Astyanax mexicanus. *J Comp Neurol,* 505**,** 221-33.

ALVAREZ-VIEJO, M., CERNUDA-CERNUDA, R., ALVAREZ-LOPEZ, C. & GARCIA-FERNANDEZ, J. M. 2004. Identification of extraretinal photoreceptors in the teleost Phoxinus phoxinus. *Histol Histopathol,* 19**,** 487-94.

AMORES, A., FORCE, A., YAN, Y. L., JOLY, L., AMEMIYA, C., FRITZ, A., HO, R. K., LANGELAND, J., PRINCE, V., WANG, Y. L., WESTERFIELD, M., EKKER, M. & POSTLETHWAIT, J. H. 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science,* 282**,** 1711-4.

ARENDT, D. 2003. Evolution of eyes and photoreceptor cell types. *Int J Dev Biol,* 47**,** 563-71.

ARNIS, S. & HOFMANN, K. P. 1993. Two different forms of metarhodopsin II: Schiff base deprotonation precedes proton uptake and signaling state. *Proc Natl Acad Sci U S A,* 90**,** 7849-53.

ASENJO, A. B., RIM, J. & OPRIAN, D. D. 1994. Molecular determinants of human red/green color discrimination. *Neuron,* 12**,** 1131-8.

ATON, S. J., HUETTNER, J. E., STRAUME, M. & HERZOG, E. D. 2006. GABA and Gi/o differentially control circadian rhythms and synchrony in clock neurons. *Proc Natl Acad Sci U S A,* 103**,** 19188-93.

BABU, K. R., DUKKIPATI, A., BIRGE, R. R. & KNOX, B. E. 2001. Regulation of phototransduction in short-wavelength cone visual pigments via the retinylidene Schiff base counterion. *Biochemistry,* 40**,** 13760-6.

BAILES, H. J., ZHUANG, L. Y. & LUCAS, R. J. 2012. Reproducible and sustained regulation of Galphas signalling using a metazoan opsin as an optogenetic tool. *PLoS One,* 7**,** e30774.

BAILEY, M. J. & CASSONE, V. M. 2004. Opsin photoisomerases in the chick retina and pineal gland: characterization, localization, and circadian regulation. *Invest Ophthalmol Vis Sci,* 45**,** 769-75.

BALCI, M. M. & AKDEMIR, R. 2011. NKX2.5 mutations and congenital heart disease: is it a marker of cardiac anomalies? *Int J Cardiol,* 147**,** e44-5.

BALSALOBRE, A., DAMIOLA, F. & SCHIBLER, U. 1998. A serum shock induces circadian gene expression in mammalian tissue culture cells. *Cell,* 93**,** 929-37.

BAUSE, E. & LEGLER, G. 1981. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis. *Biochem J,* 195**,** 639-44.

BELLINGHAM, J., CHAURASIA, S. S., MELYAN, Z., LIU, C., CAMERON, M. A., TARTTELIN, E. E., IUVONE, P. M., HANKINS, M. W., TOSINI, G. & LUCAS, R. J. 2006. Evolution of melanopsin photoreceptors: discovery and characterization of a new melanopsin in nonmammalian vertebrates. *PLoS Biol,* 4**,** e254.

BELLINGHAM, J. & FOSTER, R. G. 2002. Opsins and mammalian photoentrainment. *Cell Tissue Res,* 309**,** 57-71.

BELLINGHAM, J., MORRIS, A. G. & HUNT, D. M. 1998. The rhodopsin gene of the cuttlefish Sepia officinalis: sequence and spectral tuning. *J Exp Biol,* 201**,** 2299-306.

BELLINGHAM, J., TARTTELIN, E. E., FOSTER, R. G. & WELLS, D. J. 2003a. Structure and evolution of the teleost extraretinal rod-like opsin (errlo) and ocular rod opsin (rho) genes: is teleost rho a retrogene? *J Exp Zool B Mol Dev Evol,* 297**,** 1-10.

BELLINGHAM, J., WELLS, D. J. & FOSTER, R. G. 2003b. In silico characterisation and chromosomal localisation of human RRH (peropsin)--implications for opsin evolution. *BMC Genomics,* 4**,** 3.

BELLINGHAM, J., WHITMORE, D., PHILP, A. R., WELLS, D. J. & FOSTER, R. G. 2002. Zebrafish melanopsin: isolation, tissue localisation and phylogenetic position. *Brain Res Mol Brain Res,* 107**,** 128-36.

BENOIST, C., O'HARE, K., BREATHNACH, R. & CHAMBON, P. 1980. The ovalbumin gene-sequence of putative control regions. *Nucleic Acids Res,* 8**,** 127-42.

BENSON, D. W., SILBERBACH, G. M., KAVANAUGH-MCHUGH, A., COTTRILL, C., ZHANG, Y., RIGGS, S., SMALLS, O., JOHNSON, M. C., WATSON, M. S., SEIDMAN, J. G., SEIDMAN, C. E., PLOWDEN, J. & KUGLER, J. D. 1999. Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest,* 104**,** 1567-73.

BENTROP, J., SCHWAB, K., PAK, W. L. & PAULSEN, R. 1997. Site-directed mutagenesis of highly conserved amino acids in the first cytoplasmic loop of Drosophila Rh1 opsin blocks rhodopsin synthesis in the nascent state. *EMBO J,* 16**,** 1600-9.

BERSON, D. M. 2003. Strange vision: ganglion cells as circadian photoreceptors. *Trends Neurosci,* 26**,** 314-20.

BERSON, D. M., DUNN, F. A. & TAKAO, M. 2002. Phototransduction by retinal ganglion cells that set the circadian clock. *Science,* 295**,** 1070-3.

BHATTACHARYA, S., WU, Z., MIYAGI, M., WEST, K., JIN, Z., NAWROT, M., SAARI, J. & CRABB, J. 2002. Interactions of CRALBP With Other Visual Cycle Proteins. *Invest. Ophthalmol. Vis. Sci.,* 43**,** 4567-.

BIEHLMAIER, O., NEUHAUSS, S. C. & KOHLER, K. 2003. Synaptic plasticity and functionality at the cone terminal of the developing zebrafish retina. *J Neurobiol,* 56**,** 222-36.

BILLUPS, D., BILLUPS, B., CHALLISS, R. A. & NAHORSKI, S. R. 2006. Modulation of Gq-protein-coupled inositol trisphosphate and Ca2+ signaling by the membrane potential. *J Neurosci,* 26**,** 9983-95.

BIRGE, R. R. & KNOX, B. E. 2003. Perspectives on the counterion switch-induced photoactivation of the G protein-coupled receptor rhodopsin. *Proc Natl Acad Sci U S A,* 100**,** 9105-7.

BLACKSHAW, S. & SNYDER, S. H. 1997. Parapinopsin, a novel catfish opsin localized to the parapineal organ, defines a new gene family. *J Neurosci,* 17**,** 8083-92.

BLACKSHAW, S. & SNYDER, S. H. 1999. Encephalopsin: A novel mammalian extraretinal opsin discretely localized in the brain. *Journal of Neuroscience,* 19**,** 3681-3690.

BOROWSKY, R. 2008a. Astyanax mexicanus, the Blind Mexican Cave Fish: A Model for Studies in Development and Morphology. *CSH Protoc,* 2008**,** pdb emo107.

BOROWSKY, R. 2008b. The Extraordinary Evolution of 'Blind' Cavefish - Mexican tetra or Blind Cave Fish (*Astyanax mexicanus*).

BOVA, L. M., WOOD, A. M., JAMIE, J. F. & TRUSCOTT, R. J. 1999. UV filter compounds in human lenses: the origin of 4-(2-amino-3-hydroxyphenyl)-4-oxobutanoic acid O-beta-D-glucoside. *Invest Ophthalmol Vis Sci,* 40**,** 3237-44.

BRADIC, M., BEERLI, P., GARCIA-DE LEON, F. J., ESQUIVEL-BOBADILLA, S. & BOROWSKY, R. L. 2012. Gene flow and population structure in the Mexican blind cavefish complex (Astyanax mexicanus). *BMC Evol Biol,* 12**,** 9.

BRADSHAW, R. A. & DENNIS, E. A. 2004. *Handbook of cell signaling,* Amsterdam San Diego, Calif., Academic Press.

BRAINARD, G. C. & HANIFIN, J. P. 2005. Photons, clocks, and consciousness. *J Biol Rhythms,* 20**,** 314-25.

BRIDGER, J. M. & VOLPI, E. V. 2010. *Fluorescence in situ hybridization (FISH) : protocols and applications,* New York, NY, Humana Press.

BRIGGS, W. R. & SPUDICH, J. L. 2006. *Handbook of Photosensory Receptors*, John Wiley & Sons.

BRITT, L. L., LOEW, E. R. & MCFARLAND, W. N. 2001. Visual pigments in the early life stages of Pacific northwest marine fishes. *J Exp Biol,* 204**,** 2581-7.

BROWN, H. M., ITO, H. & OGDEN, T. E. 1968. Spectral sensitivity of the planarian ocellus. *J Gen Physiol,* 51**,** 255-60.

BROWN, R. L. & ROBINSON, P. R. 2004. Melanopsin--shedding light on the elusive circadian photopigment. *Chronobiol Int,* 21**,** 189-204.

BURKE, T. W., WILLY, P. J., KUTACH, A. K., BUTLER, J. E. & KADONAGA, J. T. 1998. The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold Spring Harb Symp Quant Biol,* 63**,** 75-82.

BURNSIDE, B. & ACKLAND, N. 1984. Effects of circadian rhythm and cAMP on retinomotor movements in the green sunfish, Lepomis cyanellus. *Invest Ophthalmol Vis Sci,* 25**,** 539-45.

CAO, P., SUN, W., KRAMP, K., ZHENG, M., SALOM, D., JASTRZEBSKA, B., JIN, H., PALCZEWSKI, K. & FENG, Z. 2012. Light-sensitive coupling of rhodopsin and melanopsin to G(i/o) and G(q) signal transduction in Caenorhabditis elegans. *FASEB J,* 26**,** 480-91.

CARROLL, J. & JACOBS, G. H. 2008. Mammalian photopigments. *In:* MASLAND, R. H. & KANEKO, A. (eds.) *The Senses: A Comprehensive Reference. Vol. 1.* Oxford: Elsevier.

CARVALHO, L. S., COWING, J. A., WILKIE, S. E., BOWMAKER, J. K. & HUNT, D. M. 2007. The molecular evolution of avian ultraviolet- and violet-sensitive visual pigments. *Mol Biol Evol,* 24**,** 1843-52.

CAVALLARI, N., FRIGATO, E., VALLONE, D., FROHLICH, N., LOPEZ-OLMEDA, J. F., FOA, A., BERTI, R., SANCHEZ-VAZQUEZ, F. J., BERTOLUCCI, C. & FOULKES, N. S. 2011. A blind circadian clock in cavefish reveals that opsins mediate peripheral clock photoreception. *PLoS Biol,* 9**,** e1001142.

CERNY, R., HORÁČEK, I. & OLSSON, L. 2006. The Trabecula cranii: development and homology of an enigmatic vertebrate head structure. *Animal Biology,* 56**,** 503-518.

CHEN, J., RATTNER, A. & NATHANS, J. 2005a. The rod photoreceptor-specific nuclear receptor Nr2e3 represses transcription of multiple cone-specific genes. *J Neurosci,* 25**,** 118-29.

CHEN, P., HAO, W., RIFE, L., WANG, X. P., SHEN, D., CHEN, J., OGDEN, T., VAN BOEMEL, G. B., WU, L., YANG, M. & FONG, H. K. 2001a. A photic visual cycle of rhodopsin regeneration is dependent on Rgr. *Nat Genet,* 28**,** 256-60.

CHEN, P., KOCHOUNIAN, H. & FONG, H. K. W. 2005b. Protein Complex Components of RGR Opsin Isolated From the Bovine Retinal Pigment Epithelium. *Invest. Ophthalmol. Vis. Sci.,* 46**,** 5127-.

CHEN, P., LEE, T. D. & FONG, H. K. 2001b. Interaction of 11-cis-retinol dehydrogenase with the chromophore of retinal g protein-coupled receptor opsin. *J Biol Chem,* 276**,** 21098-104.

CHEN, S. & ZACK, D. J. 1996. Ret 4, a positive acting rhodopsin regulatory element identified using a bovine retina in vitro transcription system. *J Biol Chem,* 271**,** 28549-57.

CHINEN, A., MATSUMOTO, Y. & KAWAMURA, S. 2005. Reconstitution of ancestral green visual pigments of zebrafish and molecular mechanism of their spectral differentiation. *Mol Biol Evol,* 22**,** 1001-10.

CHOE, H. W., PARK, J. H., KIM, Y. J. & ERNST, O. P. 2011. Transmembrane signaling by GPCRs: insight from rhodopsin and opsin structures. *Neuropharmacology,* 60**,** 52-7.

CHOW, B. Y., CHUONG, A. S., KLAPOETKE, N. C. & BOYDEN, E. S. 2011. Synthetic physiology strategies for adapting tools from nature for genetically targeted control of fast biological processes. *Methods Enzymol,* 497**,** 425-43.

COLLI, L., PAGLIANTI, A., BERTI, R., GANDOLFI, G. & TAGLIAVINI, J. 2009. Molecular phylogeny of the blind cavefish Phreatichthys andruzzii and Garra barreimiae within the family Cyprinidae. *Environmental Biology of Fishes,* 84**,** 95-107.

COLLIN, J. P., VOISIN, P., FALCON, J., FAURE, J. P., BRISSON, P. & DEFAYE, J. R. 1989. Pineal transducers in the course of evolution: molecular organization, rhythmic metabolic activity and role. *Arch Histol Cytol,* 52 Suppl**,** 441-9.

COLLIN, S. P., HART, N. S., SHAND, J. & POTTER, I. C. 2003a. Morphology and spectral absorption characteristics of retinal photoreceptors in the southern hemisphere lamprey (Geotria australis). *Vis Neurosci,* 20**,** 119-30.

COLLIN, S. P., KNIGHT, M. A., DAVIES, W. L., POTTER, I. C., HUNT, D. M. & TREZISE, A. E. 2003b. Ancient colour vision: multiple opsin genes in the ancestral vertebrates. *Current Biology,* 13**,** R864-5.

COLLIN, S. P. & MARSHALL, N. J. 2003. *Sensory processing in aquatic environments,* New York, Springer.

CONCEPCION, F., MENDEZ, A. & CHEN, J. 2002. The carboxyl-terminal domain is essential for rhodopsin transport in rod photoreceptors. *Vision Res,* 42**,** 417-26.

CONNAUGHTON, V. P., GRAHAM, D. & NELSON, R. 2004. Identification and morphological classification of horizontal, bipolar, and amacrine cells within the zebrafish retina. *J Comp Neurol,* 477**,** 371-85.

CONTI, A. C., CRYAN, J. F., DALVI, A., LUCKI, I. & BLENDY, J. A. 2002. cAMP response element-binding protein is essential for the upregulation of brain-derived neurotrophic factor transcription, but not the behavioral or endocrine responses to antidepressant drugs. *J Neurosci,* 22**,** 3262-8.

COSTA, R. H., KALINICHENKO, V. V., HOLTERMAN, A. X. & WANG, X. 2003. Transcription factors in liver development, differentiation, and regeneration. *Hepatology,* 38**,** 1331-47.

COWING, J. A., POOPALASUNDARAM, S., WILKIE, S. E., ROBINSON, P. R., BOWMAKER, J. K. & HUNT, D. M. 2002. The molecular mechanism for the spectral shifts between vertebrate ultraviolet- and violet-sensitive cone visual pigments. *Biochem J,* 367**,** 129-35.

COX, M. M. 1999. Recombinational DNA repair in bacteria and the RecA protein. *Prog Nucleic Acid Res Mol Biol,* 63**,** 311-66.

CRESCITELLI, F., MCFALL-NGAI, M. & HORWITZ, J. 1985. The visual pigment sensitivity hypothesis: further evidence from fishes of varying habitats. *J Comp Physiol A,* 157**,** 323-33.

CRONIN, T. W., SHASHAR, N., CALDWELL, R. L., MARSHALL, J., CHEROSKE, A. G. & CHIOU, T. H. 2003. Polarization vision and its role in biological signaling. *Integr Comp Biol,* 43**,** 549-58.

CROSS, S. H. & BIRD, A. P. 1995. CpG islands and genes. *Curr Opin Genet Dev,* 5**,** 309-14.

CUGINI, P., DI PALMA, L., DI SIMONE, S., LUCIA, P., BATTISTI, P., COPPOLA, A. & LEONE, G. 1993. Circadian rhythm of cardiac output, peripheral vascular resistance, and related variables by a beat-to-beat monitoring. *Chronobiol Int,* 10**,** 73-8.

CULVER, D. C. & PIPAN, T. 2009. *The biology of caves and other subterranean habitats,* Oxford, Oxford University Press.

CZEISLER, C. A. & GOOLEY, J. J. 2007. Sleep and circadian rhythms in humans. *Cold Spring Harb Symp Quant Biol,* 72**,** 579-97.

DACEY, D. M., LIAO, H. W., PETERSON, B. B., ROBINSON, F. R., SMITH, V. C., POKORNY, J., YAU, K. W. & GAMLIN, P. D. 2005. Melanopsin-expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN. *Nature,* 433**,** 749-54.

DARTNALL, H. J. & LYTHGOE, J. N. 1965. The spectral clustering of visual pigments. *Vision Res,* 5**,** 81-100.

DAVIDSON, F. F., LOEWEN, P. C. & KHORANA, H. G. 1994. Structure and function in rhodopsin: replacement by alanine of cysteine residues 110 and 187, components of a conserved disulfide

bond in rhodopsin, affects the light-activated metarhodopsin II state. *Proc Natl Acad Sci U S A,* 91**,** 4029-33.

DAVIES, W. I. 2011. Adaptive Gene Loss in Vertebrates: Photosensitivity as a Model Case. *Encyclopedia of Life Sciences***,** 1-10.

DAVIES, W. I., COLLIN, S. P. & HUNT, D. M. 2012a. Molecular ecology and adaptation of visual photopigments in craniates. *Mol Ecol,* 21**,** 3121-58.

DAVIES, W. I., DOWNES, S. M., FU, J. K., SHANKS, M. E., COPLEY, R. R., LISE, S., RAMSDEN, S. C., BLACK, G. C., GIBSON, K., FOSTER, R. G., HANKINS, M. W. & NEMETH, A. H. 2012b. Next-generation sequencing in health-care delivery: lessons from the functional analysis of rhodopsin. *Genet Med,* 14**,** 891-9.

DAVIES, W. I., TAY, B. H., ZHENG, L., DANKS, J. A., BRENNER, S., FOSTER, R. G., COLLIN, S. P., HANKINS, M. W., VENKATESH, B. & HUNT, D. M. 2012c. Evolution and Functional Characterisation of Melanopsins in a Deep-Sea Chimaera (Elephant Shark, Callorhinchus milii). *PLoS One,* 7**,** e51276.

DAVIES, W. I., TURTON, M., PEIRSON, S. N., FOLLETT, B. K., HALFORD, S., GARCIA-FERNANDEZ, J. M., SHARP, P. J., HANKINS, M. W. & FOSTER, R. G. 2012d. Vertebrate ancient opsin photopigment spectra and the avian photoperiodic response. *Biol Lett,* 8**,** 291-4.

DAVIES, W. I., WILKIE, S. E., COWING, J. A., HANKINS, M. W. & HUNT, D. M. 2012e. Anion sensitivity and spectral tuning of middle- and long-wavelength-sensitive (MWS/LWS) visual pigments. *Cell Mol Life Sci,* 69**,** 2455-64.

DAVIES, W. I., ZHENG, L., HUGHES, S., TAMAI, T. K., TURTON, M., HALFORD, S., FOSTER, R. G., WHITMORE, D. & HANKINS, M. W. 2011. Functional diversity of melanopsins and their global expression in the teleost retina. *Cell Mol Life Sci,* 68**,** 4115-32.

DAVIES, W. L., CARVALHO, L. S. & HUNT, D. M. 2007a. SPLICE: a technique for generating in vitro spliced coding sequences from genomic DNA. *Biotechniques,* 43**,** 785-9.

DAVIES, W. L., COWING, J. A., CARVALHO, L. S., POTTER, I. C., TREZISE, A. E., HUNT, D. M. & COLLIN, S. P. 2007b. Functional characterization, tuning, and regulation of visual pigment gene expression in an anadromous lamprey. *FASEB J,* 21**,** 2713-24.

DAVIES, W. L., FOSTER, R. G. & HANKINS, M. W. 2012f. Focus on molecules: melanopsin. *Exp Eye Res,* 97**,** 161-2.

DAVIES, W. L., HANKINS, M. W. & FOSTER, R. G. 2010. Vertebrate ancient opsin and melanopsin: divergent irradiance detectors. *Photochem Photobiol Sci,* 9**,** 1444-57.

DEEB, S. S. 2004. Molecular genetics of color-vision deficiencies. *Vis Neurosci,* 21**,** 191-6.

DEEB, S. S., WAKEFIELD, M. J., TADA, T., MAROTTE, L., YOKOYAMA, S. & MARSHALL GRAVES, J. A. 2003. The cone visual pigments of an Australian marsupial, the tammar wallaby (Macropus eugenii): sequence, spectral tuning, and evolution. *Mol Biol Evol,* 20**,** 1642-9.

DEININGER, P. L. & BATZER, M. A. 2002. Mammalian Retroelements. *Genome Research,* 12**,** 1455-1465.

DEKENS, M. P., SANTORIELLO, C., VALLONE, D., GRASSI, G., WHITMORE, D. & FOULKES, N. S. 2003. Light regulates the cell cycle in zebrafish. *Curr Biol,* 13**,** 2051-7.

DEKENS, M. P. & WHITMORE, D. 2008. Autonomous onset of the circadian clock in the zebrafish embryo. *EMBO J,* 27**,** 2757-65.

DI ROCCO, G., PENNUTO, M., ILLI, B., CANU, N., FILOCAMO, G., TRANI, E., RINALDI, A. M., POSSENTI, R., MANDOLESI, G., SIRINIAN, M. I., JUCKER, R., LEVI, A. & NASI, S. 1997. Interplay of the E box, the cyclic AMP response element, and HTF4/HEB in transcriptional regulation of the neurospecific, neurotrophin-inducible vgf gene. *Mol Cell Biol,* 17**,** 1244-53.

DIEL, S., KLASS, K., WITTIG, B. & KLEUSS, C. 2006. Gbetagamma activation site in adenylyl cyclase type II. Adenylyl cyclase type III is inhibited by Gbetagamma. *J Biol Chem,* 281**,** 288-94.

DIJK, D. J. & ARCHER, S. N. 2009. Light, sleep, and circadian rhythms: together again. *PLoS Biol,* 7**,** e1000145.

DOI, M., ISHIDA, A., MIYAKE, A., SATO, M., KOMATSU, R., YAMAZAKI, F., KIMURA, I., TSUCHIYA, S., KORI, H., SEO, K., YAMAGUCHI, Y., MATSUO, M., FUSTIN, J. M., TANAKA, R., SANTO, Y., YAMADA, H., TAKAHASHI, Y., ARAKI, M., NAKAO, K., AIZAWA, S., KOBAYASHI, M., OBRIETAN, K., TSUJIMOTO, G. & OKAMURA, H. 2011. Circadian regulation of intracellular G-protein signalling mediates intercellular synchrony and rhythmicity in the suprachiasmatic nucleus. *Nat Commun,* 2**,** 327.

DOI, T., MOLDAY, R. S. & KHORANA, H. G. 1990. Role of the intradiscal domain in rhodopsin assembly and function. *Proc Natl Acad Sci U S A,* 87**,** 4991-5.

DOLPH, P. J., RANGANATHAN, R., COLLEY, N. J., HARDY, R. W., SOCOLICH, M. & ZUKER, C. S. 1993. Arrestin function in inactivation of G protein-coupled receptor rhodopsin in vivo. *Science,* 260**,** 1910-6.

DORVAL, K. M., BOBECHKO, B. P., FUJIEDA, H., CHEN, S., ZACK, D. J. & BREMNER, R. 2006. CHX10 targets a subset of photoreceptor genes. *J Biol Chem,* 281**,** 744-51.

DOUGLAS, R. & DJAMGOZ, M. B. A. 1990. *The Visual system of fish,* London ; New York, Chapman and Hall.

DRYJA, T. P., BERSON, E. L., RAO, V. R. & OPRIAN, D. D. 1993. Heterozygous missense mutation in the rhodopsin gene as a cause of congenital stationary night blindness. *Nat Genet,* 4**,** 280-3.

DRYJA, T. P., MCGEE, T. L., HAHN, L. B., COWLEY, G. S., OLSSON, J. E., REICHEL, E., SANDBERG, M. A. & BERSON, E. L. 1990a. Mutations within the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa. *N Engl J Med,* 323**,** 1302-7.

DRYJA, T. P., MCGEE, T. L., REICHEL, E., HAHN, L. B., COWLEY, G. S., YANDELL, D. W., SANDBERG, M. A. & BERSON, E. L. 1990b. A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. *Nature,* 343**,** 364-6.

DUKKIPATI, A., KUSNETZOW, A., BABU, K. R., RAMOS, L., SINGH, D., KNOX, B. E. & BIRGE, R. R. 2002. Phototransduction by vertebrate ultraviolet visual pigments: protonation of the retinylidene Schiff base following photobleaching. *Biochemistry,* 41**,** 9842-51.

DUKKIPATI, A., VOUGHT, B. W., SINGH, D., BIRGE, R. R. & KNOX, B. E. 2001. Serine 85 in transmembrane helix 2 of short-wavelength visual pigments interacts with the retinylidene Schiff base counterion. *Biochemistry,* 40**,** 15098-108.

DUNCAN, S. A., MANOVA, K., CHEN, W. S., HOODLESS, P., WEINSTEIN, D. C., BACHVAROVA, R. F. & DARNELL, J. E., JR. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proc Natl Acad Sci U S A,* 91**,** 7598-602.

EMEIS, D., KUHN, H., REICHERT, J. & HOFMANN, K. P. 1982. Complex formation between metarhodopsin II and GTP-binding protein in bovine photoreceptor membranes leads to a shift of the photoproduct equilibrium. *FEBS Lett,* 143**,** 29-34.

ERCKENS, W. & MARTIN, W. 1982. Exogenous and endogenous control of swimming activity in Astyanax mexicanus (Characidae, Pisces) by direct light response and by a circadian oscillator. II. Features of time-controlled behaviour of a cave population and their comparison to an epigean ancestral form. *Verlag der Zeitschrift für Naturforschung,* 37**,** 1253-1265.

ESKIN, A., CORRENT, G., LIN, C. Y. & MCADOO, D. J. 1982. Mechanism for shifting the phase of a circadian rhythm by serotonin: involvement of cAMP. *Proc Natl Acad Sci U S A,* 79**,** 660-4.

ESPINASA, L. & JEFFERY, W. R. 2006. Conservation of retinal circadian rhythms during cavefish eye degeneration. *Evol Dev,* 8**,** 16-22.

FAIN, G. L., MATTHEWS, H. R., CORNWALL, M. C. & KOUTALOS, Y. 2001. Adaptation in vertebrate photoreceptors. *Physiol Rev,* 81**,** 117-151.

FALVEY, E., MARCACCI, L. & SCHIBLER, U. 1996. DNA-binding specificity of PAR and C/EBP leucine zipper proteins: a single amino acid substitution in the C/EBP DNA-binding domain confers PAR-like specificity to C/EBP. *Biol Chem,* 377**,** 797-809.

FARHAT, F. P., MARTINS, C. B., DE LIMA, L. H., ISOLDI, M. C. & CASTRUCCI, A. M. 2009. Melanopsin and clock genes: regulation by light and endothelin in the zebrafish ZEM-2S cell line. *Chronobiol Int,* 26**,** 1090-119.

FARRAR, G. J., FINDLAY, J. B., KUMAR-SINGH, R., KENNA, P., HUMPHRIES, M. M., SHARPE, E. & HUMPHRIES, P. 1992. Autosomal dominant retinitis pigmentosa: a novel mutation in the rhodopsin gene in the original 3q linked family. *Hum Mol Genet,* 1**,** 769-71.

FARRENS, D. L. & KHORANA, H. G. 1995. Structure and function in rhodopsin. Measurement of the rate of metarhodopsin II decay by fluorescence spectroscopy. *J Biol Chem,* 270**,** 5073-6.

FASICK, J. I., APPLEBURY, M. L. & OPRIAN, D. D. 2002. Spectral tuning in the mammalian short-wavelength sensitive cone pigments. *Biochemistry,* 41**,** 6860-5.

FISHER, M. M. & WEISS, K. 1974. Laser photolysis of retinal and its protonated and unprotonated n-butylamine Schiff base. *Photochemistry and Photobiology,* 20**,** 423-432.

FITZGIBBON, J., HOPE, A., SLOBODYANYUK, S. J., BELLINGHAM, J., BOWMAKER, J. K. & HUNT, D. M. 1995. The rhodopsin-encoding gene of bony fish lacks introns. *Gene,* 164**,** 273-7.

FOA, A., BASAGLIA, F., BELTRAMI, G., CARNACINA, M., MORETTO, E. & BERTOLUCCI, C. 2009. Orientation of lizards in a Morris water-maze: roles of the sun compass and the parietal eye. *Journal of Experimental Biology,* 212**,** 2918-24.

FORSELL, J., EKSTROM, P., FLAMARIQUE, I. N. & HOLMQVIST, B. 2001. Expression of pineal ultraviolet- and green-like opsins in the pineal organ and retina of teleosts. *J Exp Biol,* 204**,** 2517-25.

FOSTER, R. & BELLINGHAM, J. 2002. Opsins and melanopsins. *Curr Biol,* 12**,** R543-44.

FOSTER, R. G., GRACE, M. S., PROVENCIO, I., DEGRIP, W. J. & GARCIA-FERNANDEZ, J. M. 1994. Identification of vertebrate deep brain photoreceptors. *Neurosci Biobehav Rev,* 18**,** 541-6.

FOSTER, R. G. & HANKINS, M. W. 2002. Non-rod, non-cone photoreception in the vertebrates. *Prog Retin Eye Res,* 21**,** 507-27.

FOSTER, R. G., PROVENCIO, I., HUDSON, D., FISKE, S., DE GRIP, W. & MENAKER, M. 1991. Circadian photoreception in the retinally degenerate mouse (rd/rd). *J Comp Physiol A,* 169**,** 39-50.

FOSTER, R. G. & SONI, B. G. 1998. Extraretinal photoreceptors and their regulation of temporal physiology. *Rev Reprod,* 3**,** 145-50.

FOSTER, R. G., WAGNER, H. J. & BOWMAKER, J. K. 2006. Non-image-forming photoreception. *In:* LADICH, F., COLLIN, S. P., MOLLER, P. & KAPOOR, B. G. (eds.) *Communication in Fishes.* New Hampshire: Science Publisher Inc.

FRANKE, R. R., SAKMAR, T. P., GRAHAM, R. M. & KHORANA, H. G. 1992a. Structure and Function in Rhodopsin - Studies of the Interaction between the Rhodopsin Cytoplasmic Domain and Transducin. *Journal of Biological Chemistry,* 267**,** 14767-14774.

FRANKE, R. R., SAKMAR, T. P., GRAHAM, R. M. & KHORANA, H. G. 1992b. Structure and function in rhodopsin. Studies of the interaction between the rhodopsin cytoplasmic domain and transducin. *J Biol Chem,* 267**,** 14767-74.

FREEDMAN, M. S., LUCAS, R. J., SONI, B., VON SCHANTZ, M., MUNOZ, M., DAVID-GRAY, Z. & FOSTER, R. 1999. Regulation of mammalian circadian behavior by non-rod, non-cone, ocular photoreceptors. *Science,* 284**,** 502-4.

FRIGATO, E., VALLONE, D., BERTOLUCCI, C. & FOULKES, N. S. 2006. Isolation and characterization of melanopsin and pinopsin expression within photoreceptive sites of reptiles. *Naturwissenschaften,* 93**,** 379-85.

FRITZE, O., FILIPEK, S., KUKSA, V., PALCZEWSKI, K., HOFMANN, K. P. & ERNST, O. P. 2003a. Role of the conserved NPxxY(x)5,6F motif in the rhodopsin ground state and during activation. *Proc Natl Acad Sci U S A,* 100**,** 2290-5.

FRITZE, O., FILIPEK, S., KUKSA, V., PALCZEWSKI, K., HOFMANN, K. P. & ERNST, O. P. 2003b. Role of the conserved NPxxY(x)(5,6)F motif in the rhodopsin ground state and during activation. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 2290-2295.

FURUKAWA, T., MORROW, E. M. & CEPKO, C. L. 1997. Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell,* 91**,** 531-41.

FUTTERMAN, A. & FUTTERMAN, S. 1974. The stability of 11-cis-retinal and reactivity toward nucleophiles. *Biochimica Et Biophysica Acta,* 337**,** 390-394.

GAMLIN, P. D., MCDOUGAL, D. H., POKORNY, J., SMITH, V. C., YAU, K. W. & DACEY, D. M. 2007. Human and macaque pupil responses driven by melanopsin-containing retinal ganglion cells. *Vision Res,* 47**,** 946-54.

GARTNER, W. & TOWNER, P. 1995. Invertebrate visual pigments. *Photochem Photobiol,* 62**,** 1-16.

GINGRAS, A. C., RAUGHT, B. & SONENBERG, N. 1999. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem,* 68**,** 913-63.

GONZALEZ-FERNANDEZ, F. 2012. Interphotoreceptor retinoid binding protein; myths and mysteries. *J Ophthalmic Vis Res,* 7**,** 100-4.

GONZALEZ, G. A. & MONTMINY, M. R. 1989. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell,* 59**,** 675-80.

GOOLEY, J. J., HO MIEN, I., ST HILAIRE, M. A., YEO, S. C., CHUA, E. C., VAN REEN, E., HANLEY, C. J., HULL, J. T., CZEISLER, C. A. & LOCKLEY, S. W. 2012. Melanopsin and rod-cone photoreceptors play different roles in mediating pupillary light responses during exposure to continuous light in humans. *J Neurosci,* 32**,** 14242-53.

GOOLEY, J. J., LU, J., CHOU, T. C., SCAMMELL, T. E. & SAPER, C. B. 2001. Melanopsin in cells of origin of the retinohypothalamic tract. *Nat Neurosci,* 4**,** 1165.

GOVARDOVSKII, V. I., FYHRQUIST, N., REUTER, T., KUZMIN, D. G. & DONNER, K. 2000. In search of the visual pigment template. *Vis Neurosci,* 17**,** 509-28.

GRIFFITHS, A. J. F. 2000. *An introduction to genetic analysis,* New York, W.H. Freeman.

GRIFFITHS, H. R., MISTRY, P., HERBERT, K. E. & LUNEC, J. 1998. Molecular and cellular effects of ultraviolet light-induced genotoxicity. *Crit Rev Clin Lab Sci,* 35**,** 189-237.

GRINDLEY, J. C., DAVIDSON, D. R. & HILL, R. E. 1995. The role of Pax-6 in eye and nasal development. *Development,* 121**,** 1433-42.

GROSS, J. B. 2012. The complex origin of Astyanax cavefish. *BMC Evol Biol,* 12**,** 105.

GU, Y., OBERWINKLER, J., POSTMA, M. & HARDIE, R. C. 2005. Mechanisms of light adaptation in Drosophila photoreceptors. *Curr Biol,* 15**,** 1228-34.

GUNER-ATAMAN, B., PAFFETT-LUGASSY, N., ADAMS, M. S., NEVIS, K. R., JAHANGIRI, L., OBREGON, P., KIKUCHI, K., POSS, K. D., BURNS, C. E. & BURNS, C. G. 2013. Zebrafish second heart field development relies on progenitor specification in anterior lateral plate mesoderm and nkx2.5 function. *Development,* 140**,** 1353-63.

HALFORD, S., BELLINGHAM, J., OCAKA, L., FOX, M., JOHNSON, S., FOSTER, R. G. & HUNT, D. M. 2001a. Assignment of panopsin (OPN3) to human chromosome band 1q43 by in situ hybridization and somatic cell hybrids. *Cytogenet Cell Genet,* 95**,** 234-5.

HALFORD, S., FREEDMAN, M. S., BELLINGHAM, J., INGLIS, S. L., POOPALASUNDARAM, S., SONI, B. G., FOSTER, R. G. & HUNT, D. M. 2001b. Characterization of a novel human opsin gene with wide tissue expression and identification of embedded and flanking genes on chromosome 1q43. *Genomics,* 72**,** 203-8.

HALFORD, S., PIRES, S. S., TURTON, M., ZHENG, L., GONZALEZ-MENENDEZ, I., DAVIES, W. L., PEIRSON, S. N., GARCIA-FERNANDEZ, J. M., HANKINS, M. W. & FOSTER, R. G. 2009. VA opsin-based photoreceptors in the hypothalamus of birds. *Curr Biol,* 19**,** 1396-402.

HALL, M. D., HOON, M. A., RYBA, N. J., POTTINGER, J. D., KEEN, J. N., SAIBIL, H. R. & FINDLAY, J. B. 1991. Molecular cloning and primary structure of squid (Loligo forbesi) rhodopsin, a phospholipase C-directed G-protein-linked receptor. *Biochem J,* 274 (Pt 1)**,** 35-40.

HALL, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* 41**,** 95-98.

HALPIN, D. M. 2008. ABCD of the phosphodiesterase family: interaction and differential activity in COPD. *Int J Chron Obstruct Pulmon Dis,* 3**,** 543-61.

HANKINS, M. W., PEIRSON, S. N. & FOSTER, R. G. 2008. Melanopsin: an exciting photopigment. *Trends Neurosci,* 31**,** 27-36.

HANSEN, A. & ZEISKE, E. 1998. The peripheral olfactory organ of the zebrafish, Danio rerio: an ultrastructural study. *Chem Senses,* 23**,** 39-48.

HAO, W. & FONG, H. K. 1996. Blue and ultraviolet light-absorbing opsin from the retinal pigment epithelium. *Biochemistry,* 35**,** 6251-6.

HARA, T. & HARA, R. 1967. Rhodopsin and retinochrome in the squid retina. *Nature,* 214**,** 573-5.

HARA, T. & HARA, R. 1976. Distribution of rhodopsin and retinochrome in the squid retina. *J Gen Physiol,* 67**,** 791-805.

HARA, T., HARA, R. & TAKEUCHI, J. 1967. Rhodopsin and retinochrome in the octopus retina. *Nature,* 214**,** 572-3.

HARDIE, R. C. & RAGHU, P. 2001. Visual transduction in Drosophila. *Nature,* 413**,** 186-193.

HARDING, H. P. & LAZAR, M. A. 1993. The orphan receptor Rev-ErbA alpha activates transcription via a novel response element. *Mol Cell Biol,* 13**,** 3113-21.

HARGRAVE, P. A. 1977. The amino-terminal tryptic peptide of bovine rhodopsin. A glycopeptide containing two sites of oligosaccharide attachment. *Biochim Biophys Acta,* 492**,** 83-94.

HARGRAVE, P. A., MCDOWELL, J. H., CURTIS, D. R., WANG, J. K., JUSZCZAK, E., FONG, S. L., RAO, J. K. & ARGOS, P. 1983. The structure of bovine rhodopsin. *Biophys Struct Mech,* 9**,** 235-44.

HARGRAVE, P. A., MCDOWELL, J. H., FELDMANN, R. J., ATKINSON, P. H., RAO, J. K. & ARGOS, P. 1984. Rhodopsin's protein and carbohydrate structure: selected aspects. *Vision Res,* 24**,** 1487-99.

HAROSI, F. I. 1994. An analysis of two spectral properties of vertebrate visual pigments. *Vision Res,* 34**,** 1359-67.

HARTWICK, A. T., BRAMLEY, J. R., YU, J., STEVENS, K. T., ALLEN, C. N., BALDRIDGE, W. H., SOLLARS, P. J. & PICKARD, G. E. 2007. Light-evoked calcium responses of isolated melanopsin-expressing retinal ganglion cells. *Journal of Neuroscience,* 27**,** 13468-80.

HASTINGS, M. H. & HERZOG, E. D. 2004. Clock genes, oscillators, and cellular networks in the suprachiasmatic nuclei. *J Biol Rhythms,* 19**,** 400-13.

HATORI, M. & PANDA, S. 2010. The emerging roles of melanopsin in behavioral adaptation to light. *Trends Mol Med,* 16**,** 435-46.

HATTAR, S., LIAO, H. W., TAKAO, M., BERSON, D. M. & YAU, K. W. 2002. Melanopsin-containing retinal ganglion cells: architecture, projections, and intrinsic photosensitivity. *Science,* 295**,** 1065-70.

HATTAR, S., LUCAS, R. J., MROSOVSKY, N., THOMPSON, S., DOUGLAS, R. H., HANKINS, M. W., LEM, J., BIEL, M., HOFMANN, F., FOSTER, R. G. & YAU, K. W. 2003. Melanopsin and rod-cone photoreceptive systems account for all major accessory visual functions in mice. *Nature,* 424**,** 76-81.

HEASMAN, J. 2002. Morpholino oligos: making sense of antisense? *Dev Biol,* 243**,** 209-14.

HECK, M., SCHADEL, S. A., MARETZKI, D. & HOFMANN, K. P. 2003. Secondary binding sites of retinoids in opsin: characterization and role in regeneration. *Vision Res,* 43**,** 3003-10.

HIGGINS, D. G., THOMPSON, J. D. & GIBSON, T. J. 1996. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol,* 266**,** 383-402.

HILLMAN, P., HOCHSTEIN, S. & MINKE, B. 1983. Transduction in invertebrate photoreceptors: role of pigment bistability. *Physiol Rev,* 63**,** 668-772.

HIRAYAMA, J., CARDONE, L., DOI, M. & SASSONE-CORSI, P. 2005. Common pathways in circadian and cell cycle clocks: light-dependent activation of Fos/AP-1 in zebrafish controls CRY-1a and WEE-1. *Proc Natl Acad Sci U S A,* 102**,** 10194-9.

HOEGG, S., BRINKMANN, H., TAYLOR, J. S. & MEYER, A. 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *J Mol Evol,* 59**,** 190-203.

HOLLAND, L. Z., ALBALAT, R., AZUMI, K., BENITO-GUTIERREZ, E., BLOW, M. J., BRONNER-FRASER, M., BRUNET, F., BUTTS, T., CANDIANI, S., DISHAW, L. J., FERRIER, D. E., GARCIA-FERNANDEZ, J., GIBSON-BROWN, J. J., GISSI, C., GODZIK, A., HALLBOOK, F., HIROSE, D., HOSOMICHI, K., IKUTA, T., INOKO, H., KASAHARA, M., KASAMATSU, J., KAWASHIMA, T., KIMURA, A., KOBAYASHI, M., KOZMIK, Z., KUBOKAWA, K., LAUDET, V., LITMAN, G. W., MCHARDY, A. C., MEULEMANS, D., NONAKA, M., OLINSKI, R. P., PANCER, Z., PENNACCHIO, L. A., PESTARINO, M., RAST, J. P., RIGOUTSOS, I., ROBINSON-RECHAVI, M., ROCH, G., SAIGA, H., SASAKURA, Y., SATAKE, M., SATOU, Y., SCHUBERT, M., SHERWOOD, N., SHIINA, T., TAKATORI, N., TELLO, J., VOPALENSKY, P., WADA, S., XU, A., YE, Y., YOSHIDA, K., YOSHIZAKI, F., YU, J. K., ZHANG, Q., ZMASEK, C. M., DE JONG, P. J., OSOEGAWA, K., PUTNAM, N. H., ROKHSAR, D. S., SATOH, N. & HOLLAND, P. W. 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res,* 18**,** 1100-11.

HOLTHUES, H., ENGEL, L., SPESSERT, R. & VOLLRATH, L. 2005. Circadian gene expression patterns of melanopsin and pinopsin in the chick pineal gland. *Biochem Biophys Res Commun,* 326**,** 160-5.

HORSTMANN, V., HUETHER, C. M., JOST, W., RESKI, R. & DECKER, E. L. 2004. Quantitative promoter analysis in Physcomitrella patens: a set of plant vectors activating gene expression within three orders of magnitude. *BMC Biotechnol,* 4**,** 13.

HUBBARD, R. & KROPF, A. 1958. The Action of Light on Rhodopsin. *Proc Natl Acad Sci U S A,* 44**,** 130-9.

HUGHES, S., HANKINS, M. W., FOSTER, R. G. & PEIRSON, S. N. 2012a. Melanopsin phototransduction: slowly emerging from the dark. *Prog Brain Res,* 199**,** 19-40.

HUGHES, S., WELSH, L., KATTI, C., GONZALEZ-MENENDEZ, I., TURTON, M., HALFORD, S., SEKARAN, S., PEIRSON, S. N., HANKINS, M. W. & FOSTER, R. G. 2012b. Differential expression of melanopsin isoforms Opn4L and Opn4S during postnatal development of the mouse retina. *PLoS One,* 7**,** e34531.

HUNT, D. M., COWING, J. A., WILKIE, S. E., PARRY, J. W., POOPALASUNDARAM, S. & BOWMAKER, J. K. 2004. Divergent mechanisms for the tuning of shortwave sensitive visual pigments in vertebrates. *Photochem Photobiol Sci,* 3**,** 713-20.

HUNT, D. M., DULAI, K. S., COWING, J. A., JULLIOT, C., MOLLON, J. D., BOWMAKER, J. K., LI, W. H. & HEWETT-EMMETT, D. 1998. Molecular evolution of trichromacy in primates. *Vision Res,* 38**,** 3299-306.

HUNT, D. M., WILKIE, S. E., BOWMAKER, J. K. & POOPALASUNDARAM, S. 2001. Vision in the ultraviolet. *Cell Mol Life Sci,* 58**,** 1583-98.

HWANG, Y. C., ZHENG, Q., GREGORY, B. D. & WANG, L. S. 2013. High-throughput identification of long-range regulatory elements and their target promoters in the human genome. *Nucleic Acids Res*.

ILLING, M. E., RAJAN, R. S., BENCE, N. F. & KOPITO, R. R. 2002. A rhodopsin mutant linked to autosomal dominant retinitis pigmentosa is prone to aggregate and interacts with the ubiquitin proteasome system. *J Biol Chem,* 277**,** 34150-60.

IMAI, H., TERAKITA, A., TACHIBANAKI, S., IMAMOTO, Y., YOSHIZAWA, T. & SHICHIDA, Y. 1997. Photochemical and biochemical properties of chicken blue-sensitive cone visual pigment. *Biochemistry,* 36**,** 12773-9.

IMAMOTO, Y., SEKI, I., YAMASHITA, T. & SHICHIDA, Y. 2013. Efficiencies of Activation of Transducin by Cone and Rod Visual Pigments. *Biochemistry*.

ISHIHARA, K. & SASAKI, H. 2002. An evolutionarily conserved putative insulator element near the 3' boundary of the imprinted Igf2/H19 domain. *Hum Mol Genet,* 11**,** 1627-36.

ISOLDI, M. C., ROLLAG, M. D., CASTRUCCI, A. M. & PROVENCIO, I. 2005. Rhabdomeric phototransduction initiated by the vertebrate photopigment melanopsin. *Proc Natl Acad Sci U S A,* 102**,** 1217-21.

JACOBS, G. H., NEITZ, M., DEEGAN, J. F. & NEITZ, J. 1996. Trichromatic colour vision in New World monkeys. *Nature,* 382**,** 156-8.

JEFFERY, W. R. 2001. Cavefish as a model system in evolutionary developmental biology. *Dev Biol,* 231**,** 1-12.

JEFFERY, W. R. 2005. Adaptive evolution of eye degeneration in the Mexican blind cavefish. *J Hered,* 96**,** 185-96.

JEFFERY, W. R. 2009. Regressive evolution in Astyanax cavefish. *Annu Rev Genet,* 43**,** 25-47.

JERLOV, N. G. 1976. *Marine optics,* Amsterdam ; New York, Elsevier Scientific Pub. Co.

JETTEN, A. M., KUREBAYASHI, S. & UEDA, E. 2001. The ROR nuclear orphan receptor subfamily: critical regulators of multiple biological processes. *Prog Nucleic Acid Res Mol Biol,* 69**,** 205-47.

JIANG, M., PANDEY, S. & FONG, H. K. 1993. An opsin homologue in the retina and pigment epithelium. *Invest Ophthalmol Vis Sci,* 34**,** 3669-78.

JIN, M., LI, S., NUSINOWITZ, S., LLOYD, M., HU, J., RADU, R. A., BOK, D. & TRAVIS, G. H. 2009. The role of interphotoreceptor retinoid-binding protein on the translocation of visual retinoids and function of cone photoreceptors. *J Neurosci,* 29**,** 1486-95.

JONES, A. V., CAMPBELL, P. J., BEER, P. A., SCHNITTGER, S., VANNUCCHI, A. M., ZOI, K., PERCY, M. J., MCMULLIN, M. F., SCOTT, L. M., TAPPER, W., SILVER, R. T., OSCIER, D., HARRISON, C. N., GRALLERT, H., KISIALIOU, A., STRIKE, P., CHASE, A. J., GREEN, A. R. & CROSS, N. C. 2010. The JAK2 46/1 haplotype predisposes to MPL-mutated myeloproliferative neoplasms. *Blood,* 115**,** 4517-23.

JONES, D. T., TAYLOR, W. R. & THORNTON, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci,* 8**,** 275-82.

JONES, R. 2007. Let sleeping zebrafish lie: a new model for sleep studies. *PLoS Biol,* 5**,** e281.

JUO, Z. S., CHIU, T. K., LEIBERMAN, P. M., BAIKALOV, I., BERK, A. J. & DICKERSON, R. E. 1996. How proteins recognize the TATA box. *J Mol Biol,* 261**,** 239-54.

KAKITANI, H., KAKITANI, T., RODMAN, H. & HONIG, B. 1985. On the mechanism of wavelength regulation in visual pigments. *Photochem Photobiol,* 41**,** 471-9.

KANEKO, M., HERNANDEZ-BORSETTI, N. & CAHILL, G. M. 2006. Diversity of zebrafish peripheral oscillators revealed by luciferase reporting. *Proc Natl Acad Sci U S A,* 103**,** 14614-9.

KANWAL, J. S. & FINGER, T. E. 1997. Parallel medullary gustatospinal pathways in a catfish: possible neural substrates for taste-mediated food search. *J Neurosci,* 17**,** 4873-85.

KARNIK, S. S., SAKMAR, T. P., CHEN, H. B. & KHORANA, H. G. 1988. Cysteine residues 110 and 187 are essential for the formation of correct structure in bovine rhodopsin. *Proc Natl Acad Sci U S A,* 85**,** 8459-63.

KASAHARA, M. & MORISHITA, S. 2006. *Large-scale genome sequence processing,* London

Singapore ; Hackensack, NJ, Imperial College Press ;

Distributed by World Scientific.

KATAOKA, K., NODA, M. & NISHIZAWA, M. 1994. Maf nuclear oncoprotein recognizes sequences related to an AP-1 site and forms heterodimers with both Fos and Jun. *Mol Cell Biol,* 14**,** 700-12.

KATO, T., JR., TODO, T., AYAKI, H., ISHIZAKI, K., MORITA, T., MITRA, S. & IKENAGA, M. 1994. Cloning of a marsupial DNA photolyase gene and the lack of related nucleotide sequences in placental mammals. *Nucleic Acids Res,* 22**,** 4119-24.

KAWANO-YAMASHITA, E., TERAKITA, A., KOYANAGI, M., SHICHIDA, Y., OISHI, T. & TAMOTSU, S. 2007. Immunohistochemical characterization of a parapinopsin-containing photoreceptor cell involved in the ultraviolet/green discrimination in the pineal organ of the river lamprey Lethenteron japonicum. *J Exp Biol,* 210**,** 3821-9.

KEFALOV, V. J., CROUCH, R. K. & CORNWALL, M. C. 2001. Role of noncovalent binding of 11-cis-retinal to opsin in dark adaptation of rod and cone photoreceptors. *Neuron,* 29**,** 749-55.

KELLER, E. B. & NOON, W. A. 1984. Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. *Proc Natl Acad Sci U S A,* 81**,** 7417-20.

KERPPOLA, T. K. & CURRAN, T. 1994. Maf and Nrl can bind to AP-1 sites and form heterodimers with Fos and Jun. *Oncogene,* 9**,** 675-84.

KIM, Y., KWEON, J. & KIM, J. S. 2013. TALENs and ZFNs are associated with different mutation signatures. *Nat Methods,* 10**,** 185.

KIMURA, A., SINGH, D., WAWROUSEK, E. F., KIKUCHI, M., NAKAMURA, M. & SHINOHARA, T. 2000. Both PCE-1/RX and OTX/CRX interactions are necessary for photoreceptor-specific gene expression. *J Biol Chem,* 275**,** 1152-60.

KIMURA, M., MAEDA, K. & HAYASHI, S. 1992. Cytosolic calcium increase in coronary endothelial cells after H2O2 exposure and the inhibitory effect of U78517F. *Br J Pharmacol,* 107**,** 488-93.

KIRCHBERG, K., KIM, T. Y., MOLLER, M., SKEGRO, D., DASARA RAJU, G., GRANZIN, J., BULDT, G., SCHLESINGER, R. & ALEXIEV, U. 2011. Conformational dynamics of helix 8 in the GPCR rhodopsin controls arrestin activation in the desensitization process. *Proc Natl Acad Sci U S A,* 108**,** 18690-5.

KISELEV, A. & SUBRAMANIAM, S. 1997. Studies of Rh1 metarhodopsin stabilization in wild-type Drosophila and in mutants lacking one or both arrestins. *Biochemistry,* 36**,** 2188-96.

KITO, Y., SUZUKI, T., AZUMA, M. & SEKOGUTI, Y. 1968. Absorption spectrum of rhodopsin denatured with acid. *Nature,* 218**,** 955-7.

KLEINAU, G., JAESCHKE, H., WORTH, C. L., MUELLER, S., GONZALEZ, J., PASCHKE, R. & KRAUSE, G. 2010. Principles and determinants of G-protein coupling by the rhodopsin-like thyrotropin receptor. *PLoS One,* 5**,** e9745.

KLEINJAN, D. A. & VAN HEYNINGEN, V. 2005. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet,* 76**,** 8-32.

KLEINSCHMIDT, J. & HAROSI, F. I. 1992. Anion sensitivity and spectral tuning of cone visual pigments in situ. *Proc Natl Acad Sci U S A,* 89**,** 9181-5.

KOBAYASHI, Y., ISHIKAWA, T., HIRAYAMA, J., DAIYASU, H., KANAI, S., TOH, H., FUKUDA, I., TSUJIMURA, T., TERADA, N., KAMEI, Y., YUBA, S., IWAI, S. & TODO, T. 2000. Molecular analysis of zebrafish photolyase/cryptochrome family: two types of cryptochromes present in zebrafish. *Genes Cells,* 5**,** 725-38.

KOBILKA, B. K., KOBILKA, T. S., DANIEL, K., REGAN, J. W., CARON, M. G. & LEFKOWITZ, R. J. 1988. Chimeric alpha 2-,beta 2-adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. *Science,* 240**,** 1310-6.

KOJIMA, D. & FUKADA, Y. 1999. Non-visual photoreception by a variety of vertebrate opsins. *Novartis Found Symp,* 224**,** 265-79; discussion 279-82.

KOJIMA, D., MANO, H. & FUKADA, Y. 2000. Vertebrate ancient-long opsin: a green-sensitive photoreceptive molecule present in zebrafish deep brain and retinal horizontal cells. *Journal of Neuroscience,* 20**,** 2845-51.

KOJIMA, D., MORI, S., TORII, M., WADA, A., MORISHITA, R. & FUKADA, Y. 2011. UV-sensitive photoreceptor protein OPN5 in humans and mice. *PLoS One,* 6**,** e26388.

KONARSKA, M. M., GRABOWSKI, P. J., PADGETT, R. A. & SHARP, P. A. 1985. Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. *Nature,* 313**,** 552-7.

KONO, M. 2006. Constitutive activity of a UV cone opsin. *FEBS Lett,* 580**,** 229-32.

KOYANAGI, M., KAWANO, E., KINUGAWA, Y., OISHI, T., SHICHIDA, Y., TAMOTSU, S. & TERAKITA, A. 2004. Bistable UV pigment in the lamprey pineal. *Proc Natl Acad Sci U S A,* 101**,** 6687-91.

KOYANAGI, M., NAGATA, T., KATOH, K., YAMASHITA, S. & TOKUNAGA, F. 2008a. Molecular evolution of arthropod color vision deduced from multiple opsin genes of jumping spiders. *J Mol Evol,* 66**,** 130-7.

KOYANAGI, M., TAKADA, E., NAGATA, T., TSUKAMOTO, H. & TERAKITA, A. 2013. Homologs of vertebrate Opn3 potentially serve as a light sensor in nonphotoreceptive tissue. *Proc Natl Acad Sci U S A,* 110**,** 4998-5003.

KOYANAGI, M., TAKANO, K., TSUKAMOTO, H., OHTSU, K., TOKUNAGA, F. & TERAKITA, A. 2008b. Jellyfish vision starts with cAMP signaling mediated by opsin-G(s) cascade. *Proc Natl Acad Sci U S A,* 105**,** 15576-80.

KOYANAGI, M., TERAKITA, A., KUBOKAWA, K. & SHICHIDA, Y. 2002. Amphioxus homologs of Go-coupled rhodopsin and peropsin having 11-cis- and all-trans-retinals as their chromophores. *FEBS Lett,* 531**,** 525-8.

KOZAK, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res,* 15**,** 8125-48.

KOZAK, M. 1991a. An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol,* 115**,** 887-903.

KOZAK, M. 1991b. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem,* 266**,** 19867-70.

KOZMIK, Z. 2005. Pax genes in eye development and evolution. *Curr Opin Genet Dev,* 15**,** 430-8.

KRISTIANSEN, K. 2004. Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function. *Pharmacol Ther,* 103**,** 21-80.

KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E. L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol,* 305**,** 567-80.

KSANTINI, M., SENECHAL, A., HUMBERT, G., ARNAUD, B. & HAMEL, C. P. 2007. RRH, encoding the RPE-expressed opsin-like peropsin, is not mutated in retinitis pigmentosa and allied diseases. *Ophthalmic Genet,* 28**,** 31-7.

KULKARNI, M. M. 2011. Digital multiplexed gene expression analysis using the NanoString nCounter system. *Curr Protoc Mol Biol,* Chapter 25**,** Unit25B 10.

KUMBALASIRI, T., ROLLAG, M. D., ISOLDI, M. C., CASTRUCCI, A. M. & PROVENCIO, I. 2007. Melanopsin triggers the release of internal calcium stores in response to light. *Photochem Photobiol,* 83**,** 273-9.

KURAKU, S. & KURATANI, S. 2011. Genome-wide detection of gene extinction in early mammalian evolution. *Genome Biol Evol,* 3**,** 1449-62.

KUSAKABE, T., KUSAKABE, R., KAWAKAMI, I., SATOU, Y., SATOH, N. & TSUDA, M. 2001. Ci-opsin1, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian Ciona intestinalis. *FEBS Lett,* 506**,** 69-72.

KUSNETZOW, A. K., DUKKIPATI, A., BABU, K. R., RAMOS, L., KNOX, B. E. & BIRGE, R. R. 2004. Vertebrate ultraviolet visual pigments: protonation of the retinylidene Schiff base and a counterion switch during photoactivation. *Proc Natl Acad Sci U S A,* 101**,** 941-6.

LADICH, F. 2006. *Communication in fishes,* Enfield, NH, Science Publishers.

LAGRANGE, T., KAPANIDIS, A. N., TANG, H., REINBERG, D. & EBRIGHT, R. H. 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev,* 12**,** 34-44.

LALL, G. S., REVELL, V. L., MOMIJI, H., AL ENEZI, J., ALTIMUS, C. M., GULER, A. D., AGUILAR, C., CAMERON, M. A., ALLENDER, S., HANKINS, M. W. & LUCAS, R. J. 2010. Distinct contributions of rod, cone, and melanopsin photoreceptors to encoding irradiance. *Neuron,* 66**,** 417-28.

LAMB, T. D., COLLIN, S. P. & PUGH, E. N., JR. 2007. Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. *Nat Rev Neurosci,* 8**,** 960-76.

LAND, M. F. & FERNALD, R. D. 1992. The evolution of eyes. *Annu Rev Neurosci,* 15**,** 1-29.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LEE, W., MITCHELL, P. & TJIAN, R. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell,* 49**,** 741-52.

LEECH, D. M. & JOHNSEN, S. 2009. Light, Biological Receptors. *In:* LIKENS, G. E. (ed.) *Encyclopedia of Inland Waters.* Oxford: Elsevier.

LEHNINGER, A. L., NELSON, D. L. & COX, M. M. 2013. *Lehninger principles of biochemistry,* New York, W.H. Freeman.

LESSER, M. P., CARLETON, K. L., BOTTGER, S. A., BARRY, T. M. & WALKER, C. W. 2011. Sea urchin tube feet are photosensory organs that express a rhabdomeric-like opsin and PAX6. *Proc Biol Sci,* 278**,** 3371-9.

LEWIS-TUFFIN, L. J., QUINN, P. G. & CHIKARAISHI, D. M. 2004. Tyrosine hydroxylase transcription depends primarily on cAMP response element activity, regardless of the type of inducing stimulus. *Mol Cell Neurosci,* 25**,** 536-47.

LIM, C. Y., SANTOSO, B., BOULAY, T., DONG, E., OHLER, U. & KADONAGA, J. T. 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev,* 18**,** 1606-17.

LISTON, D. R. & JOHNSON, P. J. 1999. Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. *Mol Cell Biol,* 19**,** 2380-8.

LODISH, H. F. 2004. *Molecular cell biology,* New York, W.H. Freeman and Company.

LONZE, B. E. & GINTY, D. D. 2002. Function and regulation of CREB family transcription factors in the nervous system. *Neuron,* 35**,** 605-23.

LOPEZ-OLMEDA, J. F., TARTAGLIONE, E. V., DE LA IGLESIA, H. O. & SANCHEZ-VAZQUEZ, F. J. 2010. Feeding entrainment of food-anticipatory activity and per1 expression in the brain and liver of zebrafish under different lighting and feeding conditions. *Chronobiol Int,* 27**,** 1380-400.

LUCAS-LLEDO, J. I. & LYNCH, M. 2009. Evolution of mutation rates: phylogenomic analysis of the photolyase/cryptochrome family. *Mol Biol Evol,* 26**,** 1143-53.

LUCAS, R. J. 2006. Chromophore regeneration: melanopsin does its own thing. *Proc Natl Acad Sci U S A,* 103**,** 10153-4.

LUCAS, R. J., DOUGLAS, R. H. & FOSTER, R. G. 2001. Characterization of an ocular photopigment capable of driving pupillary constriction in mice. *Nat Neurosci,* 4**,** 621-6.

LUCAS, R. J., FREEDMAN, M. S., MUNOZ, M., GARCIA-FERNANDEZ, J. M. & FOSTER, R. G. 1999. Regulation of the mammalian pineal by non-rod, non-cone, ocular photoreceptors. *Science,* 284**,** 505-7.

LUTCKE, H. A., CHOW, K. C., MICKEL, F. S., MOSS, K. A., KERN, H. F. & SCHEELE, G. A. 1987. Selection of AUG initiation codons differs in plants and animals. *EMBO J,* 6**,** 43-8.

LYUBARSKY, A. L., FALSINI, B., PENNESI, M. E., VALENTINI, P. & PUGH, E. N., JR. 1999. UV- and midwave-sensitive cone-driven retinal responses of the mouse: a possible phenotype for coexpression of cone photopigments. *J Neurosci,* 19**,** 442-55.

MAEDA, T., VAN HOOSER, J. P., DRIESSEN, C. A., FILIPEK, S., JANSSEN, J. J. & PALCZEWSKI, K. 2003. Evaluation of the role of the retinal G protein-coupled receptor (RGR) in the vertebrate retina in vivo. *J Neurochem,* 85**,** 944-56.

MANNA, P. R. & STOCCO, D. M. 2007. Crosstalk of CREB and Fos/Jun on a single cis-element: transcriptional repression of the steroidogenic acute regulatory protein gene. *J Mol Endocrinol,* 39**,** 261-77.

MANO, H., KOJIMA, D. & FUKADA, Y. 1999. Exo-rhodopsin: a novel rhodopsin expressed in the zebrafish pineal gland. *Brain Res Mol Brain Res,* 73**,** 110-8.

MASON, B., SCHMALE, M., GIBBS, P., MILLER, M. W., WANG, Q., LEVAY, K., SHESTOPALOV, V. & SLEPAK, V. Z. 2012. Evidence for multiple phototransduction pathways in a reef-building coral. *PLoS One,* 7**,** e50371.

MASSARI, M. E. & MURRE, C. 2000. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol,* 20**,** 429-40.

MAX, M., SURYA, A., TAKAHASHI, J. S., MARGOLSKEE, R. F. & KNOX, B. E. 1998. Light-dependent activation of rod transducin by pineal opsin. *J Biol Chem,* 273**,** 26820-6.

MCCLEMENTS, M., DAVIES, W. I., MICHAELIDES, M., YOUNG, T., NEITZ, M., MACLAREN, R. E., MOORE, A. T. & HUNT, D. M. 2013. Variations in opsin coding sequences cause x-linked cone dysfunction syndrome with myopia and dichromacy. *Invest Ophthalmol Vis Sci,* 54**,** 1361-9.

MCFARLAND, W. N. & ALLEN, D. M. 1977. The effect of extrinsic factors on two distinctive rhodopsin-porphyropsin systems. *Can J Zool,* 55**,** 1000-9.

MEJIA, R. 2011. Cave-dwelling fish provide clues to the circadian cycle. *PLoS Biol,* 9**,** e1001141.

MELCHIOR, B. & FRANGOS, J. A. 2012. Galphaq/11-mediated intracellular calcium responses to retrograde flow in endothelial cells. *Am J Physiol Cell Physiol,* 303**,** C467-73.

MELYAN, Z., TARTTELIN, E. E., BELLINGHAM, J., LUCAS, R. J. & HANKINS, M. W. 2005. Addition of human melanopsin renders mammalian cells photoresponsive. *Nature,* 433**,** 741-5.

MENDEZ, F. & PENNER, R. 1998. Near-visible ultraviolet light induces a novel ubiquitous calcium-permeable cation current in mammalian cell lines. *J Physiol,* 507 (Pt 2)**,** 365-77.

MENGER, G. J., KOKE, J. R. & CAHILL, G. M. 2005. Diurnal and circadian retinomotor movements in zebrafish. *Vis Neurosci,* 22**,** 203-9.

MENON, S. T., HAN, M. & SAKMAR, T. P. 2001. Rhodopsin: structural basis of molecular physiology. *Physiol Rev,* 81**,** 1659-88.

MEYER, A. & SCHARTL, M. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol,* 11**,** 699-704.

MINAMOTO, T. & SHIMIZU, I. 2002. A novel isoform of vertebrate ancient opsin in a smelt fish, Plecoglossus altivelis. *Biochem Biophys Res Commun,* 290**,** 280-6.

MITCHELL, R. W., RUSSELL, W. H. & ELLIOTT, W. R. 1977. *Mexican eyeless characin fishes, genus Astyanax : environment, distribution, and evolution,* Lubbock, Texas Tech Press.

MIYAGUCHI, K., KUO, C. H., MIKI, N. & HASHIMOTO, P. H. 1992. Topography of Opsin within Disk and Plasma-Membranes Revealed by a Rapid-Freeze Deep-Etch Technique. *Journal of Neurocytology,* 21**,** 807-819.

MOENS, C. B. & PRINCE, V. E. 2002. Constructing the hindbrain: insights from the zebrafish. *Dev Dyn,* 224**,** 1-17.

MOLDAY, R. S. & MACKENZIE, D. 1983. Monoclonal antibodies to rhodopsin: characterization, cross-reactivity, and application as structural probes. *Biochemistry,* 22**,** 653-60.

MONTMINY, M. R., SEVARINO, K. A., WAGNER, J. A., MANDEL, G. & GOODMAN, R. H. 1986. Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. *Proc Natl Acad Sci U S A,* 83**,** 6682-6.

MORIMURA, H., SAINDELLE-RIBEAUDEAU, F., BERSON, E. L. & DRYJA, T. P. 1999. Mutations in RGR, encoding a light-sensitive opsin homologue, in patients with retinitis pigmentosa. *Nat Genet,* 23**,** 393-4.

MORROW, B. E. 2006. Microdeletions and Microduplications: Mechanism. *Encyclopedia of Life Sciences.* 27 JAN 2006 ed.: John Wiley & Sons, Ltd.

MORROW, J. M., LAZIC, S. & CHANG, B. S. 2011. A novel rhodopsin-like gene expressed in zebrafish retina. *Vis Neurosci,* 28**,** 325-35.

MOSELEY, H. 1988. *Non-ionising radiation : microwaves, ultraviolet, and laser radiation,* Bristol ; Philadelphia, A. Hilger, in collaboration with the Hospital Physicists' Association.

MOSES, K. & RUBIN, G. M. 1991. Glass encodes a site-specific DNA-binding protein that is regulated in response to positional signals in the developing Drosophila eye. *Genes Dev,* 5**,** 583-93.

MOUNT, D. W. 2007. Using the Basic Local Alignment Search Tool (BLAST). *CSH Protoc,* 2007**,** pdb top17.

MOUNT, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res,* 10**,** 459-72.

MOUTSAKI, P., BELLINGHAM, J., SONI, B. G., DAVID-GRAY, Z. K. & FOSTER, R. G. 2000. Sequence, genomic structure and tissue expression of carp (Cyprinus carpio L.) vertebrate ancient (VA) opsin. *FEBS Lett,* 473**,** 316-22.

MOUTSAKI, P., WHITMORE, D., BELLINGHAM, J., SAKAMOTO, K., DAVID-GRAY, Z. K. & FOSTER, R. G. 2003. Teleost multiple tissue (tmt) opsin: a candidate photopigment regulating the peripheral clocks of zebrafish? *Brain Res Mol Brain Res,* 112**,** 135-45.

MUELLER, K. P. & NEUHAUSS, S. C. 2012. Light perception: more than meets the eyes. *Curr Biol,* 22**,** R912-4.

MUNOZ, E. & BALER, R. 2003. The circadian E-box: when perfect is not good enough. *Chronobiol Int,* 20**,** 371-88.

MURE, L. S., CORNUT, P. L., RIEUX, C., DROUYER, E., DENIS, P., GRONFIER, C. & COOPER, H. M. 2009. Melanopsin bistability: a fly's eye technology in the human retina. *PLoS One,* 4**,** e5991.

MURE, L. S., RIEUX, C., HATTAR, S. & COOPER, H. M. 2007. Melanopsin-dependent nonvisual responses: Evidence for photopigment bistability in vivo. *Journal of Biological Rhythms,* 22**,** 411-424.

NAKANE, Y., IKEGAMI, K., ONO, H., YAMAMOTO, N., YOSHIDA, S., HIRUNAGI, K., EBIHARA, S., KUBO, Y. & YOSHIMURA, T. 2010. A mammalian neural tissue opsin (Opsin 5) is a deep brain photoreceptor in birds. *Proc Natl Acad Sci U S A,* 107**,** 15264-8.

NATHANS, J. 1990a. Determinants of Visual Pigment Absorbance - Role of Charged Amino-Acids in the Putative Transmembrane Segments. *Biochemistry,* 29**,** 937-942.

NATHANS, J. 1990b. Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. *Biochemistry,* 29**,** 9746-52.

NATHANS, J. & HOGNESS, D. S. 1983. Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell,* 34**,** 807-14.

NATHANS, J., PIANTANIDA, T. P., EDDY, R. L., SHOWS, T. B. & HOGNESS, D. S. 1986. Molecular genetics of inherited variation in human color vision. *Science,* 232**,** 203-10.

NIKOLAOU, N. & MEYER, M. P. 2012. Imaging circuit formation in zebrafish. *Dev Neurobiol,* 72**,** 346-57.

NISHIMURA, T., OKANO, H., TADA, H., NISHIMURA, E., SUGIMOTO, K., MOHRI, K. & FUKUSHIMA, M. 2010. Lizards respond to an extremely low-frequency electromagnetic field. *Journal of Experimental Biology,* 213**,** 1985-90.

NOONAN, J. P. & MCCALLION, A. S. 2010. Genomics of long-range regulatory elements. *Annu Rev Genomics Hum Genet,* 11**,** 1-23.

NOREN, M. 2003. *Danio rerio*.

O'NEILL, J. S., MAYWOOD, E. S., CHESHAM, J. E., TAKAHASHI, J. S. & HASTINGS, M. H. 2008. cAMP-dependent signaling as a core component of the mammalian circadian pacemaker. *Science,* 320**,** 949-53.

ODEEN, A. & HASTAD, O. 2003. Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. *Mol Biol Evol,* 20**,** 855-61.

ODEEN, A. & HASTAD, O. 2009. New primers for the avian SWS1 pigment opsin gene reveal new amino acid configurations in spectral sensitivity tuning sites. *J Hered,* 100**,** 784-9.

ODEEN, A., HASTAD, O. & ALSTROM, P. 2011. Evolution of ultraviolet vision in the largest avian radiation - the passerines. *BMC Evol Biol,* 11**,** 313.

OKA, Y., SARAIVA, L. R., KWAN, Y. Y. & KORSCHING, S. I. 2009. The fifth class of Galpha proteins. *Proc Natl Acad Sci U S A,* 106**,** 1484-9.

OKAJIMA, T. I., PEPPERBERG, D. R., RIPPS, H., WIGGERT, B. & CHADER, G. J. 1990. Interphotoreceptor retinoid-binding protein promotes rhodopsin regeneration in toad photoreceptors. *Proc Natl Acad Sci U S A,* 87**,** 6907-11.

OKANO, T., TAKANAKA, Y., NAKAMURA, A., HIRUNAGI, K., ADACHI, A., EBIHARA, S. & FUKADA, Y. 1997. Immunocytochemical identification of pinopsin in pineal glands of chicken and pigeon. *Brain Res Mol Brain Res,* 50**,** 190-6.

OKANO, T., YOSHIZAWA, T. & FUKADA, Y. 1994. Pinopsin is a chicken pineal photoreceptive molecule. *Nature,* 372**,** 94-7.

OVCHINNIKOV YU, A., ABDULAEV, N. G. & BOGACHUK, A. S. 1988a. Two adjacent cysteine residues in the C-terminal cytoplasmic fragment of bovine rhodopsin are palmitylated. *FEBS Lett,* 230**,** 1-5.

OVCHINNIKOV YU, A., ABDULAEV, N. G., ZOLOTAREV, A. S., ARTAMONOV, I. D., BESPALOV, I. A., DERGACHEV, A. E. & TSUDA, M. 1988b. Octopus rhodopsin. Amino acid sequence deduced from cDNA. *FEBS Lett,* 232**,** 69-72.

OZAKI, K., NAGATANI, H., OZAKI, M. & TOKUNAGA, F. 1993. Maturation of major Drosophila rhodopsin, ninaE, requires chromophore 3-hydroxyretinal. *Neuron,* 10**,** 1113-9.

PALACIOS, A. G., BOZINOVIC, F., VIELMA, A., ARRESE, C. A., HUNT, D. M. & PEICHL, L. 2010. Retinal photoreceptor arrangement, SWS1 and LWS opsin sequence, and electroretinography in the South American marsupial Thylamys elegans (Waterhouse, 1839). *J Comp Neurol,* 518**,** 1589-602.

PALCZEWSKI, K., BUCZYLKO, J., LEBIODA, L., CRABB, J. W. & POLANS, A. S. 1993. Identification of the N-terminal region in rhodopsin kinase involved in its interaction with rhodopsin. *J Biol Chem,* 268**,** 6004-13.

PALCZEWSKI, K., KUMASAKA, T., HORI, T., BEHNKE, C. A., MOTOSHIMA, H., FOX, B. A., LE TRONG, I., TELLER, D. C., OKADA, T., STENKAMP, R. E., YAMAMOTO, M. & MIYANO, M. 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science,* 289**,** 739-45.

PANDA, S., NAYAK, S. K., CAMPO, B., WALKER, J. R., HOGENESCH, J. B. & JEGLA, T. 2005. Illumination of the melanopsin signaling pathway. *Science,* 307**,** 600-4.

PANDA, S., PROVENCIO, I., TU, D. C., PIRES, S. S., ROLLAG, M. D., CASTRUCCI, A. M., PLETCHER, M. T., SATO, T. K., WILTSHIRE, T., ANDAHAZY, M., KAY, S. A., VAN GELDER, R. N. & HOGENESCH, J. B. 2003. Melanopsin is required for non-image-forming photic responses in blind mice. *Science,* 301**,** 525-7.

PANDA, S., SATO, T. K., CASTRUCCI, A. M., ROLLAG, M. D., DEGRIP, W. J., HOGENESCH, J. B., PROVENCIO, I. & KAY, S. A. 2002. Melanopsin (Opn4) requirement for normal light-induced circadian phase shifting. *Science,* 298**,** 2213-6.

PANDEY, S., BLANKS, J. C., SPEE, C., JIANG, M. & FONG, H. K. 1994. Cytoplasmic retinal localization of an evolutionary homolog of the visual pigments. *Exp Eye Res,* 58**,** 605-13.

PANDO, M. P., PINCHAK, A. B., CERMAKIAN, N. & SASSONE-CORSI, P. 2001. A cell-based system that recapitulates the dynamic light-dependent regulation of the vertebrate clock. *Proc Natl Acad Sci U S A,* 98**,** 10178-83.

PARK, J. G., PARK, Y. J., SUGAMA, N., KIM, S. J. & TAKEMURA, A. 2007. Molecular cloning and daily variations of the Period gene in a reef fish Siganus guttatus. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol,* 193**,** 403-11.

PARRY, J. W., PEIRSON, S. N., WILKENS, H. & BOWMAKER, J. K. 2003. Multiple photopigments from the Mexican blind cavefish, Astyanax fasciatus: a microspectrophotometric study. *Vision Res,* 43**,** 31-41.

PARRY, J. W., POOPALASUNDARAM, S., BOWMAKER, J. K. & HUNT, D. M. 2004. A novel amino acid substitution is responsible for spectral tuning in a rodent violet-sensitive visual pigment. *Biochemistry,* 43**,** 8014-20.

PEIRSON, S. N., HALFORD, S. & FOSTER, R. G. 2009. The evolution of irradiance detection: melanopsin and the non-visual opsins. *Philos Trans R Soc Lond B Biol Sci,* 364**,** 2849-65.

PEIRSON, S. N., OSTER, H., JONES, S. L., LEITGES, M., HANKINS, M. W. & FOSTER, R. G. 2007. Microarray analysis and functional genomics identify novel components of melanopsin signaling. *Curr Biol,* 17**,** 1363-72.

PEPE, I. M. 1999. Rhodopsin and phototransduction. *Journal of Photochemistry and Photobiology B-Biology,* 48**,** 1-10.

PEPPERBERG, D. R., BROWN, P. K., LURIE, M. & DOWLING, J. E. 1978. Visual pigment and photoreceptor sensitivity in the isolated skate retina. *J Gen Physiol,* 71**,** 369-96.

PESOLE, G., GRILLO, G., LARIZZA, A. & LIUNI, S. 2000. The untranslated regions of eukaryotic mRNAs: structure, function, evolution and bioinformatic tools for their analysis. *Brief Bioinform,* 1**,** 236-49.

PHILLIPS, J. E. & CORCES, V. G. 2009. CTCF: master weaver of the genome. *Cell,* 137**,** 1194-211.

PHILLIPS, K. & LUISI, B. 2000. The virtuoso of versatility: POU proteins that flex to fit. *J Mol Biol,* 302**,** 1023-39.

PHILP, A. R., BELLINGHAM, J., GARCIA-FERNANDEZ, J. & FOSTER, R. G. 2000a. A novel rod-like opsin isolated from the extra-retinal photoreceptors of teleost fish. *FEBS Lett,* 468**,** 181-8.

PHILP, A. R., GARCIA-FERNANDEZ, J. M., SONI, B. G., LUCAS, R. J., BELLINGHAM, J. & FOSTER, R. G. 2000b. Vertebrate ancient (VA) opsin and extraretinal photoreception in the Atlantic salmon (Salmo salar). *J Exp Biol,* 203**,** 1925-36.

PHILP, A. R., GARCIA-FERNANDEZ, J. M., SONI, B. G., LUCAS, R. J., BELLINGHAM, J. & FOSTER, R. G. 2000c. Vertebrate ancient (VA) opsin and extraretinal photoreception in the Atlantic salmon (Salmo salar). *Journal of Experimental Biology,* 203**,** 1925-36.

PIRES, S. S., HUGHES, S., TURTON, M., MELYAN, Z., PEIRSON, S. N., ZHENG, L., KOSMAOGLOU, M., BELLINGHAM, J., CHEETHAM, M. E., LUCAS, R. J., FOSTER, R. G., HANKINS, M. W. & HALFORD, S. 2009. Differential expression of two distinct functional isoforms of melanopsin (Opn4) in the mammalian retina. *J Neurosci,* 29**,** 12332-42.

PIRES, S. S., SHAND, J., BELLINGHAM, J., ARRESE, C., TURTON, M., PEIRSON, S., FOSTER, R. G. & HALFORD, S. 2007. Isolation and characterization of melanopsin (Opn4) from the Australian marsupial Sminthopsis crassicaudata (fat-tailed dunnart). *Proc Biol Sci,* 274**,** 2791-9.

PITT, G. A., COLLINS, F. D., MORTON, R. A. & STOK, P. 1955. Studies on rhodopsin. VIII. Retinylidenemethylamine, an indicator yellow analogue. *Biochem J,* 59**,** 122-8.

POINTER, M. A., CARVALHO, L. S., COWING, J. A., BOWMAKER, J. K. & HUNT, D. M. 2007. The visual pigments of a deep-sea teleost, the pearl eye Scopelarchus analis. *J Exp Biol,* 210**,** 2829-35.

PORTER, M., DITTMAR, K. & PEREZ-LOSADA, M. 2007a. How long does evolution of the troglomorphic form take? Estimating divergence times in *Astyanax mexicanus*. *Acta Carsologica,* 36**,** 173-182.

PORTER, M. L., CRONIN, T. W., MCCLELLAN, D. A. & CRANDALL, K. A. 2007b. Molecular characterization of crustacean visual pigments and the evolution of pancrustacean opsins. *Mol Biol Evol,* 24**,** 253-68.

POWELL, E. N., KLINCK, J. M., HOFMANN, E. E. & MCMANUS, M. A. 2003. Influence of water allocation and freshwater inflow on oyster production: a hydrodynamic-oyster population model for Galveston Bay, Texas, USA. *Environ Manage,* 31**,** 100-21.

PROSSER, R. A. & GILLETTE, M. U. 1989. The mammalian circadian clock in the suprachiasmatic nuclei is reset in vitro by cAMP. *J Neurosci,* 9**,** 1073-81.

PROTAS, M., TABANSKY, I., CONRAD, M., GROSS, J. B., VIDAL, O., TABIN, C. J. & BOROWSKY, R. 2008. Multi-trait evolution in a cave fish, Astyanax mexicanus. *Evol Dev,* 10**,** 196-209.

PROVENCIO, I., JIANG, G., DE GRIP, W. J., HAYES, W. P. & ROLLAG, M. D. 1998. Melanopsin: An opsin in melanophores, brain, and eye. *Proc Natl Acad Sci U S A,* 95**,** 340-5.

PROVENCIO, I., RODRIGUEZ, I. R., JIANG, G., HAYES, W. P., MOREIRA, E. F. & ROLLAG, M. D. 2000. A novel human opsin in the inner retina. *J Neurosci,* 20**,** 600-5.

QIU, X., KUMBALASIRI, T., CARLSON, S. M., WONG, K. Y., KRISHNA, V., PROVENCIO, I. & BERSON, D. M. 2005. Induction of photosensitivity by heterologous expression of melanopsin. *Nature,* 433**,** 745-9.

QTAISHAT, N. M., WIGGERT, B. & PEPPERBERG, D. R. 2005. Interphotoreceptor retinoid-binding protein (IRBP) promotes the release of all-trans retinol from the isolated retina following rhodopsin bleaching illumination. *Exp Eye Res,* 81**,** 455-63.

RADU, R. A., HU, J., PENG, J., BOK, D., MATA, N. L. & TRAVIS, G. H. 2008. Retinal pigment epithelium-retinal G protein receptor-opsin mediates light-dependent translocation of all-trans-retinyl esters for synthesis of visual chromophore in retinal pigment epithelial cells. *J Biol Chem,* 283**,** 19730-8.

RAJAMANI, R., LIN, Y. L. & GAO, J. 2011. The opsin shift and mechanism of spectral tuning in rhodopsin. *J Comput Chem,* 32**,** 854-65.

RAO, S., CHUN, C., FAN, J., KOFRON, J. M., YANG, M. B., HEGDE, R. S., FERRARA, N., COPENHAGEN, D. R. & LANG, R. A. 2013. A direct and melanopsin-dependent fetal light response regulates mouse eye development. *Nature,* 494**,** 243-6.

RAO, V. N., HUEBNER, K., ISOBE, M., AR-RUSHDI, A., CROCE, C. M. & REDDY, E. S. 1989. elk, tissue-specific ets-related genes on chromosomes X and 14 near translocation breakpoints. *Science,* 244**,** 66-70.

RAYMOND, P. A. & BARTHEL, L. K. 2004. A moving wave patterns the cone photoreceptor mosaic array in the zebrafish retina. *Int J Dev Biol,* 48**,** 935-45.

RECKEL, F., MELZER, R. R., PARRY, J. W. & BOWMAKER, J. K. 2002. The retina of five atherinomorph teleosts: photoreceptors, patterns and spectral sensitivities. *Brain Behav Evol,* 60**,** 249-64.

REHEMTULLA, A., WARWAR, R., KUMAR, R., JI, X., ZACK, D. J. & SWAROOP, A. 1996. The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. *Proc Natl Acad Sci U S A,* 93**,** 191-5.

RIDGE, K. D. & PALCZEWSKI, K. 2007. Visual rhodopsin sees the light: structure and mechanism of G protein signaling. *Journal of Biological Chemistry,* 282**,** 9297-301.

RIVOLTA, C., BERSON, E. L. & DRYJA, T. P. 2006. Mutation screening of the peropsin gene, a retinal pigment epithelium specific rhodopsin homolog, in patients with retinitis pigmentosa and allied diseases. *Mol Vis,* 12**,** 1511-5.

ROECKLEIN, K. A., ROHAN, K. J., DUNCAN, W. C., ROLLAG, M. D., ROSENTHAL, N. E., LIPSKY, R. H. & PROVENCIO, I. 2009. A missense variant (P10L) of the melanopsin (OPN4) gene in seasonal affective disorder. *J Affect Disord,* 114**,** 279-85.

ROLLAG, M. D., BERSON, D. M. & PROVENCIO, I. 2003. Melanopsin, ganglion-cell photoreceptors, and mammalian photoentrainment. *J Biol Rhythms,* 18**,** 227-34.

RZYMSKA-GRALA, I., PALCZEWSKI, P., BLAZ, M., ZMORZYNSKI, M., GOLEBIOWSKI, M. & WANYURA, H. 2012. A peculiar blow-out fracture of the inferior orbital wall complicated by extensive subcutaneous emphysema: A case report and review of the literature. *Pol J Radiol,* 77**,** 64-8.

SAITOU, N. & NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol,* 4**,** 406-25.

SAKAI, K., IMAMOTO, Y., SU, C. Y., TSUKAMOTO, H., YAMASHITA, T., TERAKITA, A., YAU, K. W. & SHICHIDA, Y. 2012. Photochemical nature of parietopsin. *Biochemistry,* 51**,** 1933-41.

SAKAMOTO, K., KADOTA, K. & OISHI, K. 2004. Light-induced phase-shifting of the peripheral circadian oscillator in the hearts of food-deprived mice. *Exp Anim,* 53**,** 471-4.

SAKAMOTO, T. & KHORANA, H. G. 1995. Structure and function in rhodopsin: the fate of opsin formed upon the decay of light-activated metarhodopsin II in vitro. *Proc Natl Acad Sci U S A,* 92**,** 249-53.

SAKMAR, T. P. 1998. Rhodopsin: a prototypical G protein-coupled receptor. *Prog Nucleic Acid Res Mol Biol,* 59**,** 1-34.

SAKMAR, T. P. 2002. Structure of rhodopsin and the superfamily of seven-helical receptors: the same and not the same. *Current Opinion in Cell Biology,* 14**,** 189-195.

SAKMAR, T. P., FRANKE, R. R. & KHORANA, H. G. 1989. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc Natl Acad Sci U S A,* 86**,** 8309-13.

SAKMAR, T. P., MENON, S. T., MARIN, E. P. & AWAD, E. S. 2002. Rhodopsin: insights from recent structural studies. *Annu Rev Biophys Biomol Struct,* 31**,** 443-84.

SALIBA, R. S., MUNRO, P. M., LUTHERT, P. J. & CHEETHAM, M. E. 2002. The cellular fate of mutant rhodopsin: quality control, degradation and aggresome formation. *J Cell Sci,* 115**,** 2907-18.

SANCAR, A. & REARDON, J. T. 2004. Nucleotide excision repair in E. coli and man. *Adv Protein Chem,* 69**,** 43-71.

SANTILLO, S., ORLANDO, P., DE PETROCELLIS, L., CRISTINO, L., GUGLIELMOTTI, V. & MUSIO, C. 2006. Evolving visual pigments: hints from the opsin-based proteins in a phylogenetically old "eyeless" invertebrate. *Biosystems,* 86**,** 3-17.

SAUGIER-VEBER, P., GOLDENBERG, A., DROUIN-GARRAUD, V., DE LA ROCHEBROCHARD, C., LAYET, V., DROUOT, N., LE MEUR, N., GILBERT-DU-SSARDIER, B., JOLY-HELAS, G., MOIROT, H., ROSSI, A., TOSI, M. & FREBOURG, T. 2006. Simple detection of genomic microdeletions and microduplications using QMPSF in patients with idiopathic mental retardation. *Eur J Hum Genet,* 14**,** 1009-17.

SCHEERER, P., PARK, J. H., HILDEBRAND, P. W., KIM, Y. J., KRAUSS, N., CHOE, H. W., HOFMANN, K. P. & ERNST, O. P. 2008. Crystal structure of opsin in its G-protein-interacting conformation. *Nature,* 455**,** 497-502.

SCHERTLER, G. F. 1998. Structure of rhodopsin. *Eye (Lond),* 12 (Pt 3b)**,** 504-10.

SCHLENKRICH, T., FLEISCHMANN, P. & HADER, D. P. 1995. Biochemical and spectroscopic characterization of the putative photoreceptor for phototaxis in amoebae of the cellular slime mould Dictyostelium discoideum. *J Photochem Photobiol B,* 30**,** 139-43.

SCHMIDT, T. M., DO, M. T., DACEY, D., LUCAS, R., HATTAR, S. & MATYNIA, A. 2011. Melanopsin-positive intrinsically photosensitive retinal ganglion cells: from form to function. *J Neurosci,* 31**,** 16094-101.

SCHMITT, E. A. & DOWLING, J. E. 1999. Early retinal development in the zebrafish, Danio rerio: light and electron microscopic analyses. *J Comp Neurol,* 404**,** 515-36.

SCHUSTER, A., WEISSCHUH, N., JAGLE, H., BESCH, D., JANECKE, A. R., ZIERLER, H., TIPPMANN, S., ZRENNER, E. & WISSINGER, B. 2005. Novel rhodopsin mutations and genotype-phenotype correlation in patients with autosomal dominant retinitis pigmentosa. *Br J Ophthalmol,* 89**,** 1258-64.

SCHUTT, F., DAVIES, S., KOPITZ, J., HOLZ, F. G. & BOULTON, M. E. 2000. Photodamage to human RPE cells by A2-E, a retinoid component of lipofuscin. *Invest Ophthalmol Vis Sci,* 41**,** 2303-8.

SEIDOU, M., SUGAHARA, M., UCHIYAMA, H., HIRAKI, K., HAMANAKA, T., MICHINOMAE, M., YOSHIHARA, K. & KITO, Y. 1990. On the three visual pigments in the retina of the firefly squid, Watasenia scintillans. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology,* 166**,** 769-773.

SEKHARAN, S., ALTUN, A. & MOROKUMA, K. 2010. Photochemistry of visual pigment in a G(q) protein-coupled receptor (GPCR)--insights from structural and spectral tuning studies on squid rhodopsin. *Chemistry,* 16**,** 1744-9.

SEKHARAN, S., SUGIHARA, M., WEINGART, O., OKADA, T. & BUSS, V. 2007. Protein assistance in the photoisomerization of rhodopsin and 9-cis-rhodopsin--insights from experiment and theory. *J Am Chem Soc,* 129**,** 1052-4.

SEKI, T. & VOGT, K. 1998. Evolutionary Aspects of the Diversity of Visual Pigment Chromophores in the Class Insecta. *Comparative Biochemistry and Physiology* 119**,** 53-64.

SEN, R., SINGH, A., K,, BALOGH-NAIR, V. & NAKANISHI, K. 1984. Photoaffinity labeling of bovine rhodopsin. *Tetrahedron,* 40**,** 493–500.

SEXTON, T., BUHR, E. & VAN GELDER, R. N. 2012. Melanopsin and mechanisms of non-visual ocular photoreception. *J Biol Chem,* 287**,** 1649-56.

SGAMBATO, V., VANHOUTTE, P., PAGES, C., ROGARD, M., HIPSKIND, R., BESSON, M. J. & CABOCHE, J. 1998. In vivo expression and regulation of Elk-1, a target of the extracellular-regulated kinase signaling pathway, in the adult rat brain. *J Neurosci,* 18**,** 214-26.

SHEN, D., JIANG, M., HAO, W., TAO, L., SALAZAR, M. & FONG, H. K. 1994. A human opsin-related gene that encodes a retinaldehyde-binding protein. *Biochemistry,* 33**,** 13117-25.

SHI, Y., RADLWIMMER, F. B. & YOKOYAMA, S. 2001. Molecular genetics and the evolution of ultraviolet vision in vertebrates. *Proc Natl Acad Sci U S A,* 98**,** 11731-6.

SHI, Y. & YOKOYAMA, S. 2003. Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates. *Proc Natl Acad Sci U S A,* 100**,** 8308-13.

SHICHIDA, Y. & IMAI, H. 1998. Visual pigment: G-protein-coupled receptor for light signals. *Cell Mol Life Sci,* 54**,** 1299-315.

SHICHIDA, Y. & MATSUYAMA, T. 2009a. Evolution of opsins and phototransduction. *Philos Trans R Soc Lond B Biol Sci,* 364**,** 2881-95.

SHICHIDA, Y. & MATSUYAMA, T. 2009b. Evolution of opsins and phototransduction. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 364**,** 2881-2895.

SHIRAKI, T., KOJIMA, D. & FUKADA, Y. 2010. Light-induced body color change in developing zebrafish. *Photochem Photobiol Sci,* 9**,** 1498-504.

SIEVING, P. A., FOWLER, M. L., BUSH, R. A., MACHIDA, S., CALVERT, P. D., GREEN, D. G., MAKINO, C. L. & MCHENRY, C. L. 2001. Constitutive "light" adaptation in rods from G90D rhodopsin: a mechanism for human congenital nightblindness without rod cell loss. *J Neurosci,* 21**,** 5449-60.

SMALL, K. M., MCGRAW, D. W. & LIGGETT, S. B. 2003. Pharmacology and physiology of human adrenergic receptor polymorphisms. *Annu Rev Pharmacol Toxicol,* 43**,** 381-411.

SMITH, K. C. 2004. Recombinational DNA repair: the ignored repair systems. *Bioessays,* 26**,** 1322-6.

SOGAWA, K., IMATAKA, H., YAMASAKI, Y., KUSUME, H., ABE, H. & FUJII-KURIYAMA, Y. 1993. cDNA cloning and transcriptional properties of a novel GC box-binding protein, BTEB2. *Nucleic Acids Res,* 21**,** 1527-32.

SOLESSIO, E. & ENGBRETSON, G. A. 1993. Antagonistic chromatic mechanisms in photoreceptors of the parietal eye of lizards. *Nature,* 364**,** 442-5.

SOLLARS, P. J., SMERASKI, C. A., KAUFMAN, J. D., OGILVIE, M. D., PROVENCIO, I. & PICKARD, G. E. 2003. Melanopsin and non-melanopsin expressing retinal ganglion cells innervate the hypothalamic suprachiasmatic nucleus. *Vis Neurosci,* 20**,** 601-10.

SONI, B. G. & FOSTER, R. G. 1997. A novel and ancient vertebrate opsin. *FEBS Lett,* 406**,** 279-83.

SONI, B. G., PHILP, A. R., FOSTER, R. G. & KNOX, B. E. 1998. Novel retinal photoreceptors. *Nature,* 394**,** 27-8.

SPARROW, J. R., PARISH, C. A., HASHIMOTO, M. & NAKANISHI, K. 1999. A2E, a lipofuscin fluorophore, in human retinal pigmented epithelial cells in culture. *Invest Ophthalmol Vis Sci,* 40**,** 2988-95.

SPENCE, R., GERLACH, G., LAWRENCE, C. & SMITH, C. 2008. The behaviour and ecology of the zebrafish, Danio rerio. *Biol Rev Camb Philos Soc,* 83**,** 13-34.

SPENCER, R. J., JIN, W., THAYER, S. A., CHAKRABARTI, S., LAW, P. Y. & LOH, H. H. 1997. Mobilization of Ca2+ from intracellular stores in transfected neuro2a cells by activation of multiple opioid receptor subtypes. *Biochem Pharmacol,* 54**,** 809-18.

SPERLING, L. & HUBBARD, R. 1975. Squid retinochrome. *J Gen Physiol,* 65**,** 235-51.

STARACE, D. M. & KNOX, B. E. 1997. Activation of transducin by a Xenopus short wavelength visual pigment. *J Biol Chem,* 272**,** 1095-100.

STARACE, D. M. & KNOX, B. E. 1998. Cloning and expression of a Xenopus short wavelength cone pigment. *Exp Eye Res,* 67**,** 209-20.

STOECKENIUS, W., LOZIER, R. H. & BOGOMOLNI, R. A. 1979. Bacteriorhodopsin and the purple membrane of halobacteria. *Biochim Biophys Acta,* 505**,** 215-78.

STOKKAN, K. A., YAMAZAKI, S., TEI, H., SAKAKI, Y. & MENAKER, M. 2001. Entrainment of the circadian clock in the liver by feeding. *Science,* 291**,** 490-3.

STRECKER, U., BERNATCHEZ, L. & WILKENS, H. 2003. Genetic divergence between cave and surface populations of Astyanax in Mexico (Characidae, Teleostei). *Mol Ecol,* 12**,** 699-710.

SU, C. Y., LUO, D. G., TERAKITA, A., SHICHIDA, Y., LIAO, H. W., KAZMI, M. A., SAKMAR, T. P. & YAU, K. W. 2006. Parietal-eye phototransduction components and their potential evolutionary implications. *Science,* 311**,** 1617-21.

SUGA, H., SCHMID, V. & GEHRING, W. J. 2008. Evolution and functional diversity of jellyfish opsins. *Curr Biol,* 18**,** 51-5.

SUN, H., GILBERT, D. J., COPELAND, N. G., JENKINS, N. A. & NATHANS, J. 1997. Peropsin, a novel visual pigment-like protein located in the apical microvilli of the retinal pigment epithelium. *Proc Natl Acad Sci U S A,* 94**,** 9893-8.

SUN, H. & NATHANS, J. 2001. ABCR, the ATP-binding cassette transporter responsible for Stargardt macular dystrophy, is an efficient target of all-trans-retinal-mediated photooxidative damage in vitro. Implications for retinal disease. *J Biol Chem,* 276**,** 11766-74.

SUNDARALINGAM, M. & BEDDELL, C. 1972. Structures of the visual chromophores and related pigments: a conformational basis of visual excitation. *Proc Natl Acad Sci U S A,* 69**,** 1569-73.

SWAROOP, A., KIM, D. & FORREST, D. 2010. Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. *Nat Rev Neurosci,* 11**,** 563-76.

TABOR, R., YAKSI, E., WEISLOGEL, J. M. & FRIEDRICH, R. W. 2004. Processing of odor mixtures in the zebrafish olfactory bulb. *J Neurosci,* 24**,** 6611-20.

TABUCHI, A., SAKAYA, H., KISUKEDA, T., FUSHIKI, H. & TSUDA, M. 2002. Involvement of an upstream stimulatory factor as well as cAMP-responsive element-binding protein in the activation of brain-derived neurotrophic factor gene promoter I. *J Biol Chem,* 277**,** 35920-31.

TAKAHASHI, J. S., DECOURSEY, P. J., BAUMAN, L. & MENAKER, M. 1984. Spectral sensitivity of a novel photoreceptive system mediating entrainment of mammalian circadian rhythms. *Nature,* 308**,** 186-8.

TAKAHASHI, J. S., MURAKAMI, N., NIKAIDO, S. S., PRATT, B. L. & ROBERTSON, L. M. 1989. The avian pineal, a vertebrate model system of the circadian oscillator: cellular regulation of circadian rhythms by light, second messengers, and macromolecular synthesis. *Recent Prog Horm Res,* 45**,** 279-348; discussion 348-52.

TAKASHIMA, N., FUJIOKA, A., HAYASAKA, N., MATSUO, A., TAKASAKI, J. & SHIGEYOSHI, Y. 2006. Gq/11-induced intracellular calcium mobilization mediates Per2 acute induction in Rat-1 fibroblasts. *Genes Cells,* 11**,** 1039-49.

TAMAI, T. K., CARR, A. J. & WHITMORE, D. 2005. Zebrafish circadian clocks: cells that see light. *Biochem Soc Trans,* 33**,** 962-6.

TAMAI, T. K., VARDHANABHUTI, V., FOULKES, N. S. & WHITMORE, D. 2004. Early embryonic light detection improves survival. *Curr Biol,* 14**,** R104-5.

TAMAI, T. K., YOUNG, L. C. & WHITMORE, D. 2007. Light signaling to the zebrafish circadian clock by Cryptochrome 1a. *Proc Natl Acad Sci U S A,* 104**,** 14712-7.

TAMURA, K., DUDLEY, J., NEI, M. & KUMAR, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol,* 24**,** 1596-9.

TAMURA, K. & NEI, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol,* 10**,** 512-26.

TANG, W. J. & GILMAN, A. G. 1991. Type-specific regulation of adenylyl cyclase by G protein beta gamma subunits. *Science,* 254**,** 1500-3.

TARTTELIN, E. E., BELLINGHAM, J., HANKINS, M. W., FOSTER, R. G. & LUCAS, R. J. 2003. Neuropsin (Opn5): a novel opsin identified in mammalian neural tissue. *FEBS Lett,* 554**,** 410-6.

TARTTELIN, E. E., FRANSEN, M. P., EDWARDS, P. C., HANKINS, M. W., SCHERTLER, G. F., VOGEL, R., LUCAS, R. J. & BELLINGHAM, J. 2011. Adaptation of pineal expressed teleost exo-rod opsin to non-image forming photoreception through enhanced Meta II decay. *Cell Mol Life Sci,* 68**,** 3713-23.

TAUSSIG, R., INIGUEZ-LLUHI, J. A. & GILMAN, A. G. 1993. Inhibition of adenylyl cyclase by Gi alpha. *Science,* 261**,** 218-21.

TAYLOR, J. S., BRAASCH, I., FRICKEY, T., MEYER, A. & VAN DE PEER, Y. 2003. Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res,* 13**,** 382-90.

TAYLOR, J. S., VAN DE PEER, Y., BRAASCH, I. & MEYER, A. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos Trans R Soc Lond B Biol Sci,* 356**,** 1661-79.

TERAKITA, A. 2005. The opsins. *Genome Biol,* 6**,** 213.

TERAKITA, A., HARA, R. & HARA, T. 1989. Retinal-binding protein as a shuttle for retinal in the rhodopsin-retinochrome system of the squid visual cells. *Vision Res,* 29**,** 639-52.

TERAKITA, A., KOYANAGI, M., TSUKAMOTO, H., YAMASHITA, T., MIYATA, T. & SHICHIDA, Y. 2004. Counterion displacement in the molecular evolution of the rhodopsin family. *Nat Struct Mol Biol,* 11**,** 284-9.

TERAKITA, A., YAMASHITA, T., NIMBARI, N., KOJIMA, D. & SHICHIDA, Y. 2002. Functional interaction between bovine rhodopsin and G protein transducin. *J Biol Chem,* 277**,** 40-6.

TERAKITA, A., YAMASHITA, T. & SHICHIDA, Y. 2000. Highly conserved glutamic acid in the extracellular IV-V loop in rhodopsins acts as the counterion in retinochrome, a member of the rhodopsin family. *Proc Natl Acad Sci U S A,* 97**,** 14263-7.

THISSE, C. & THISSE, B. 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc,* 3**,** 59-69.

THOENEN, H. 1995. Neurotrophins and neuronal plasticity. *Science,* 270**,** 593-8.

TODO, T., TAKEMORI, H., RYO, H., IHARA, M., MATSUNAGA, T., NIKAIDO, O., SATO, K. & NOMURA, T. 1993. A new photoreactivating enzyme that specifically repairs ultraviolet light-induced (6-4)photoproducts. *Nature,* 361**,** 371-4.

TOMONARI, S., MIGITA, K., TAKAGI, A., NOJI, S. & OHUCHI, H. 2008. Expression patterns of the opsin 5-related genes in the developing chicken retina. *Dev Dyn,* 237**,** 1910-22.

TOMONARI, S., TAKAGI, A., AKAMATSU, S., NOJI, S. & OHUCHI, H. 2005. A non-canonical photopigment, melanopsin, is expressed in the differentiating ganglion, horizontal, and bipolar cells of the chicken retina. *Dev Dyn,* 234**,** 783-90.

TORII, M., KOJIMA, D., OKANO, T., NAKAMURA, A., TERAKITA, A., SHICHIDA, Y., WADA, A. & FUKADA, Y. 2007. Two isoforms of chicken melanopsins show blue light sensitivity. *FEBS Lett,* 581**,** 5327-31.

TOSINI, G. & MENAKER, M. 1996. Circadian rhythms in cultured mammalian retina. *Science,* 272**,** 419-21.

TOWNSON, S. M., CHANG, B. S., SALCEDO, E., CHADWELL, L. V., PIERCE, N. E. & BRITT, S. G. 1998. Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in Drosophila, and physiological characterization. *J Neurosci,* 18**,** 2412-22.

TOYAMA, M., HIRONAKA, M., YAMAHAMA, Y., HORIGUCHI, H., TSUKADA, O., UTO, N., UENO, Y., TOKUNAGA, F., SENO, K. & HARIYAMA, T. 2008. Presence of rhodopsin and porphyropsin in the eyes of 164 fishes, representing marine, diadromous, coastal and freshwater species--a qualitative and comparative study. *Photochem Photobiol,* 84**,** 996-1002.

TRAJANO, E., BICHUETTE, M. E. & KAPOOR, B. G. 2010. *Biology of subterranean fishes,* Enfield, N.H. Boca Raton, Science Publishers ; CRC Press.

TSUKAMOTO, H., FARRENS, D. L., KOYANAGI, M. & TERAKITA, A. 2009. The magnitude of the light-induced conformational change in different rhodopsins correlates with their ability to activate G proteins. *J Biol Chem,* 284**,** 20676-83.

TSUKAMOTO, H. & TERAKITA, A. 2010. Diversity and functional properties of bistable pigments. *Photochem Photobiol Sci,* 9**,** 1435-43.

UDAGAWA, N., CHAN, J., WADA, S., FINDLAY, D. M., HAMILTON, J. A. & MARTIN, T. J. 1996. c-fos antisense DNA inhibits proliferation of osteoclast progenitors in osteoclast development but not macrophage differentiation in vitro. *Bone,* 18**,** 511-6.

UEDA, H. R., HAYASHI, S., CHEN, W., SANO, M., MACHIDA, M., SHIGEYOSHI, Y., IINO, M. & HASHIMOTO, S. 2005. System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nat Genet,* 37**,** 187-92.

URNOV, F. D., REBAR, E. J., HOLMES, M. C., ZHANG, H. S. & GREGORY, P. D. 2010. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet,* 11**,** 636-46.

VAN HOOSER, J. P., ALEMAN, T. S., HE, Y. G., CIDECIYAN, A. V., KUKSA, V., PITTLER, S. J., STONE, E. M., JACOBSON, S. G. & PALCZEWSKI, K. 2000. Rapid restoration of visual pigment and function with oral retinoid in a mouse model of childhood blindness. *Proc Natl Acad Sci U S A,* 97**,** 8623-8.

VAN OOSTERHOUT, F., FISHER, S. P., VAN DIEPEN, H. C., WATSON, T. S., HOUBEN, T., VANDERLEEST, H. T., THOMPSON, S., PEIRSON, S. N., FOSTER, R. G. & MEIJER, J. H. 2012. Ultraviolet light provides a major input to non-image-forming light detection in mice. *Curr Biol,* 22**,** 1397-402.

VATINE, G., VALLONE, D., APPELBAUM, L., MRACEK, P., BEN-MOSHE, Z., LAHIRI, K., GOTHILF, Y. & FOULKES, N. S. 2009. Light directs zebrafish period2 expression via conserved D and E boxes. *PLoS Biol,* 7**,** e1000223.

VATINE, G., VALLONE, D., GOTHILF, Y. & FOULKES, N. S. 2011. It's time to swim! Zebrafish and the circadian clock. *FEBS Lett,* 585**,** 1485-94.

VELARDE, R. A., SAUER, C. D., WALDEN, K. K., FAHRBACH, S. E. & ROBERTSON, H. M. 2005. Pteropsin: a vertebrate-like non-visual opsin expressed in the honey bee brain. *Insect Biochem Mol Biol,* 35**,** 1367-77.

VERRA, D. M., CONTIN, M. A., HICKS, D. & GUIDO, M. E. 2011. Early onset and differential temporospatial expression of melanopsin isoforms in the developing chicken retina. *Invest Ophthalmol Vis Sci,* 52**,** 5111-20.

VIGH-TEICHMANN, I., KORF, H. W., NURNBERGER, F., OKSCHE, A., VIGH, B. & OLSSON, R. 1983. Opsin-immunoreactive outer segments in the pineal and parapineal organs of the lamprey (Lampetra fluviatilis), the eel (Anguilla anguilla), and the rainbow trout (Salmo gairdneri). *Cell Tissue Res,* 230**,** 289-307.

VIGH-TEICHMANN, I., KORF, H. W., OKSCHE, A. & VIGH, B. 1982. Opsin-immunoreactive outer segments and acetylcholinesterase-positive neurons in the pineal complex of Phoxinus phoxinus (Teleostei, Cyprinidae). *Cell Tissue Res,* 227**,** 351-69.

VIGH, B., MANZANO, M. J., ZADORI, A., FRANK, C. L., LUKATS, A., ROHLICH, P., SZEL, A. & DAVID, C. 2002. Nonvisual photoreceptors of the deep brain, pineal organs and retina. *Histol Histopathol,* 17**,** 555-90.

VIGH, B., VIGH-TEICHMANN, I., ROHLICH, P. & OKSCHE, A. 1983. Cerebrospinal fluid-contacting neurons, sensory pinealocytes and Landolt's clubs of the retina as revealed by means of an electron-microscopic immunoreaction against opsin. *Cell Tissue Res,* 233**,** 539-48.

VISHNIVETSKIY, S. A., RAMAN, D., WEI, J., KENNEDY, M. J., HURLEY, J. B. & GUREVICH, V. V. 2007a. Regulation of arrestin binding by rhodopsin phosphorylation level. *Journal of Biological Chemistry,* 282**,** 32075-83.

VISHNIVETSKIY, S. A., RAMAN, D., WEI, J., KENNEDY, M. J., HURLEY, J. B. & GUREVICH, V. V. 2007b. Regulation of arrestin binding by rhodopsin phosphorylation level. *J Biol Chem,* 282**,** 32075-83.

VOGT, K. 1989. Distribution of insect visual chromophores: functional and phylogenetic aspects. *In:* STAVENGA, D. G. & HARDIE, R. C. (eds.) *Facets of vision.* Berlin: Springer.

VOGT, P. K., BOS, T. J., MITSUNOBU, F., NISHIMURA, T., MONTECLARO, F. S. & SU, H. Y. 1989. The oncogenicity of jun. *Princess Takamatsu Symp,* 20**,** 127-34.

VOPALENSKY, P., PERGNER, J., LIEGERTOVA, M., BENITO-GUTIERREZ, E., ARENDT, D. & KOZMIK, Z. 2012. Molecular analysis of the amphioxus frontal eye unravels the evolutionary origin of the retina and pigment cells of the vertebrate eye. *Proc Natl Acad Sci U S A,* 109**,** 15383-8.

WADA, S., KAWANO-YAMASHITA, E., KOYANAGI, M. & TERAKITA, A. 2012. Expression of UV-sensitive parapinopsin in the iguana parietal eyes and its implication in UV-sensitivity in vertebrate pineal-related organs. *PLoS One,* 7**,** e39003.

WAGNER, H. J. 2001. Brain areas in abyssal demersal fishes. *Brain Behav Evol,* 57**,** 301-16.

WALD, G. 1939. The Porphyropsin Visual System. *J Gen Physiol,* 22**,** 775-94.

WALD, G. 1958. The significance of vertebrate metamorphosis. *Science,* 128**,** 1481-90.

WALD, G. 1968. The molecular basis of visual excitation. *Nature,* 219**,** 800-7.

WALKER, M. T., BROWN, R. L., CRONIN, T. W. & ROBINSON, P. R. 2008. Photochemistry of retinal chromophore in mouse melanopsin. *Proc Natl Acad Sci U S A,* 105**,** 8861-5.

WALTHER, C. & GRUSS, P. 1991. Pax-6, a murine paired box gene, is expressed in the developing CNS. *Development,* 113**,** 1435-49.

WANG, H. 2008a. Comparative analysis of period genes in teleost fish genomes. *J Mol Evol,* 67**,** 29-40.

WANG, H. 2008b. Comparative analysis of teleost fish genomes reveals preservation of different ancient clock duplicates in different fishes. *Mar Genomics,* 1**,** 69-78.

WANG, H. 2009. Comparative genomic analysis of teleost fish bmal genes. *Genetica,* 136**,** 149-61.

WANG, J. & ZHOU, T. 2010. cAMP-regulated dynamics of the mammalian circadian clock. *Biosystems,* 101**,** 136-43.

WANG, Z., ASENJO, A. B. & OPRIAN, D. D. 1993. Identification of the Cl(-)-binding site in the human red and green color vision pigments. *Biochemistry,* 32**,** 2125-30.

WARRANT, E. & NILSSON, D.-E. 2006. *Invertebrate vision,* Cambridge, UK ; New York, Cambridge University Press.

WAWERSIK, S. & MAAS, R. L. 2000. Vertebrate eye development as modeled in Drosophila. *Hum Mol Genet,* 9**,** 917-25.

WEGER, B. D., SAHINBAS, M., OTTO, G. W., MRACEK, P., ARMANT, O., DOLLE, D., LAHIRI, K., VALLONE, D., ETTWILLER, L., GEISLER, R., FOULKES, N. S. & DICKMEIS, T. 2011. The light responsive transcriptome of the zebrafish: function and regulation. *PLoS One,* 6**,** e17080.

WEITZ, C. J., MIYAKE, Y., SHINZATO, K., MONTAG, E., ZRENNER, E., WENT, L. N. & NATHANS, J. 1992a. Human tritanopia associated with two amino acid substitutions in the blue-sensitive opsin. *Am J Hum Genet,* 50**,** 498-507.

WEITZ, C. J., WENT, L. N. & NATHANS, J. 1992b. Human tritanopia associated with a third amino acid substitution in the blue-sensitive visual pigment. *Am J Hum Genet,* 51**,** 444-6.

WENZEL, A., OBERHAUSER, V., PUGH, E. N., JR., LAMB, T. D., GRIMM, C., SAMARDZIJA, M., FAHL, E., SEELIGER, M. W., REME, C. E. & VON LINTIG, J. 2005. The retinal G protein-coupled receptor (RGR) enhances isomerohydrolase activity independent of light. *J Biol Chem,* 280**,** 29874-84.

WHITE, J. H., CHIANO, M., WIGGLESWORTH, M., GESKE, R., RILEY, J., WHITE, N., HALL, S., ZHU, G., MAURIO, F., SAVAGE, T., ANDERSON, W., CORDY, J., DUCCESCHI, M., VESTBO, J. & PILLAI, S. G. 2008. Identification of a novel asthma susceptibility gene on chromosome 1qter and its functional evaluation. *Hum Mol Genet,* 17**,** 1890-903.

WHITMORE, D., FOULKES, N. S. & SASSONE-CORSI, P. 2000. Light acts directly on organs and cells in culture to set the vertebrate circadian clock. *Nature,* 404**,** 87-91.

WHITMORE, D., FOULKES, N. S., STRAHLE, U. & SASSONE-CORSI, P. 1998. Zebrafish Clock rhythmic expression reveals independent peripheral circadian oscillators. *Nat Neurosci,* 1**,** 701-7.

WILKIE, S. E., ROBINSON, P. R., CRONIN, T. W., POOPALASUNDARAM, S., BOWMAKER, J. K. & HUNT, D. M. 2000. Spectral tuning of avian violet- and ultraviolet-sensitive visual pigments. *Biochemistry,* 39**,** 7895-901.

WILSON, C. J. & APPLEBURY, M. L. 1993. Arresting G-protein coupled receptor activity. *Curr Biol,* 3**,** 683-6.

WOLF, G. 1998. Transport of retinoids by the interphotoreceptor retinoid-binding protein. *Nutr Rev,* 56**,** 156-8.

WOLKEN, J. J. & MOGUS, M. A. 1979. Extra-Ocular Photosensitivity. *Photochemistry and Photobiology,* 29**,** 189-196.

WONG, K. Y. 2012. A retinal ganglion cell that can signal irradiance continuously for 10 hours. *J Neurosci,* 32**,** 11478-85.

WOOD, A. M. & TRUSCOTT, R. J. 1993. UV filters in human lenses: tryptophan catabolism. *Exp Eye Res,* 56**,** 317-25.

WOOD, P., PARTRIDGE, J. C. & DE GRIP, W. J. 1992. Rod visual pigment changes in the elver of the eel Anguilla anguilla L. measured by microspectrophotometry. *Journal ofFish Biology,* 41**,** 601-611.

WU, T., NI, Y., DONG, Y., XU, J., SONG, X., KATO, H. & FU, Z. 2010. Regulation of circadian gene expression in the kidney by light and food cues in rats. *Am J Physiol Regul Integr Comp Physiol,* 298**,** R635-41.

WULLIMANN, M. F., RUPP, B. & REICHERT, H. 1996. *Neuroanatomy of the zebrafish brain : a topological atlas,* Basel ; Boston, Birkhäuser Verlag.

XU, Y. Z., KANAGARATHAM, C., JANCIK, S. & RADZIOCH, D. 2013. Promoter deletion analysis using a dual-luciferase reporter system. *Methods Mol Biol,* 977**,** 79-93.

YAMASHITA, T., OHUCHI, H., TOMONARI, S., IKEDA, K., SAKAI, K. & SHICHIDA, Y. 2010. Opn5 is a UV-sensitive bistable pigment that couples with Gi subtype of G protein. *Proc Natl Acad Sci U S A,* 107**,** 22084-9.

YAMASHITA, T., TERAKITA, A. & SHICHIDA, Y. 2000. Distinct roles of the second and third cytoplasmic loops of bovine rhodopsin in G protein activation. *J Biol Chem,* 275**,** 34272-9.

YAMAZAKI, S., NUMANO, R., ABE, M., HIDA, A., TAKAHASHI, R., UEDA, M., BLOCK, G. D., SAKAKI, Y., MENAKER, M. & TEI, H. 2000. Resetting central and peripheral circadian oscillators in transgenic rats. *Science,* 288**,** 682-5.

YAN, E. C., KAZMI, M. A., DE, S., CHANG, B. S., SEIBERT, C., MARIN, E. P., MATHIES, R. A. & SAKMAR, T. P. 2002. Function of extracellular loop 2 in rhodopsin: glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II. *Biochemistry,* 41**,** 3620-7.

YANG-FENG, T. L., XUE, F. Y., ZHONG, W. W., COTECCHIA, S., FRIELLE, T., CARON, M. G., LEFKOWITZ, R. J. & FRANCKE, U. 1990. Chromosomal organization of adrenergic receptor genes. *Proc Natl Acad Sci U S A,* 87**,** 1516-20.

YAU, K. W. & HARDIE, R. C. 2009. Phototransduction motifs and variations. *Cell,* 139**,** 246-64.

YOKOYAMA, R. & YOKOYAMA, S. 1990. Isolation, DNA sequence and evolution of a color visual pigment gene of the blind cave fish Astyanax fasciatus. *Vision Res,* 30**,** 807-16.

YOKOYAMA, S. 1997. Molecular genetic basis of adaptive selection: examples from color vision in vertebrates. *Annu Rev Genet,* 31**,** 315-36.

YOKOYAMA, S. 2000. Molecular evolution of vertebrate visual pigments. *Prog Retin Eye Res,* 19**,** 385-419.

YOKOYAMA, S. 2008. Evolution of dim-light and color vision pigments. *Annu Rev Genomics Hum Genet,* 9**,** 259-82.

YOKOYAMA, S., RADLWIMMER, F. B. & BLOW, N. S. 2000. Ultraviolet pigments in birds evolved from violet pigments by a single amino acid change. *Proc Natl Acad Sci U S A,* 97**,** 7366-71.

YOKOYAMA, S. & SHI, Y. 2000. Genetics and evolution of ultraviolet vision in vertebrates. *FEBS Lett,* 486**,** 167-72.

YOKOYAMA, S., YANG, H. & STARMERT, W. T. 2008. Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. *Genetics,* 179**,** 2037-2043.

YOKOYAMA, S. & ZHANG, H. 1997. Cloning and characterization of the pineal gland-specific opsin gene of marine lamprey (Petromyzon marinus). *Gene,* 202**,** 89-93.

YOO, S. H., YAMAZAKI, S., LOWREY, P. L., SHIMOMURA, K., KO, C. H., BUHR, E. D., SIEPKA, S. M., HONG, H. K., OH, W. J., YOO, O. J., MENAKER, M. & TAKAHASHI, J. S. 2004. PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc Natl Acad Sci U S A,* 101**,** 5339-46.

YOSHIZAWA, M. 2008. Same Species, Different Appearance.

YOSHIZAWA, M. & JEFFERY, W. R. 2008. Shadow response in the blind cavefish Astyanax reveals conservation of a functional pineal eye. *J Exp Biol,* 211**,** 292-9.

ZELE, A. J., FEIGL, B., SMITH, S. S. & MARKWELL, E. L. 2011. The circadian response of intrinsically photosensitive retinal ganglion cells. *PLoS One,* 6**,** e17860.

ZHANG, H., FUTAMI, K., HORIE, N., OKAMURA, A., UTOH, T., MIKAWA, N., YAMADA, Y., TANAKA, S. & OKAMOTO, N. 2000. Molecular cloning of fresh water and deep-sea rod opsin genes from Japanese eel Anguilla japonica and expressional analyses during sexual maturation. *FEBS Lett,* 469**,** 39-43.

ZHANG, J., DONG, X., FUJIMOTO, Y. & OKAMURA, H. 2004. Molecular signals of Mammalian circadian clock. *Kobe J Med Sci,* 50**,** 101-9.

ZHANG, J., ZHANG, D., MCQUADE, J. S., BEHBEHANI, M., TSIEN, J. Z. & XU, M. 2002. c-fos regulates neuronal excitability and survival. *Nat Genet,* 30**,** 416-20.

ZHONG, M., KAWAGUCHI, R., KASSAI, M. & SUN, H. 2012. Retina, retinol, retinal and the natural history of vitamin a as a light sensor. *Nutrients,* 4**,** 2069-96.

ZHONG, M. & MOLDAY, R. S. 2010. Binding of retinoids to ABCA4, the photoreceptor ABC transporter associated with Stargardt macular degeneration. *Methods Mol Biol,* 652**,** 163-76.

ZHU, L., IMANISHI, Y., FILIPEK, S., ALEKSEEV, A., JASTRZEBSKA, B., SUN, W., SAPERSTEIN, D. A. & PALCZEWSKI, K. 2006. Autosomal recessive retinitis pigmentosa and E150K mutation in the opsin gene. *J Biol Chem,* 281**,** 22289-98.

ZHU, X. & CRAFT, C. M. 2000. Modulation of CRX transactivation activity by phosducin isoforms. *Mol Cell Biol,* 20**,** 5216-26.

ZHU, Y., TU, D. C., DENNER, D., SHANE, T., FITZGERALD, C. M. & VAN GELDER, R. N. 2007. Melanopsin-dependent persistence and photopotentiation of murine pupillary light responses. *Invest Ophthalmol Vis Sci,* 48**,** 1268-75.

ZHUKOVSKY, E. A. & OPRIAN, D. D. 1989. Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science,* 246**,** 928-30.

ZVONIC, S., PTITSYN, A. A., CONRAD, S. A., SCOTT, L. K., FLOYD, Z. E., KILROY, G., WU, X., GOH, B. C., MYNATT, R. L. & GIMBLE, J. M. 2006. Characterization of peripheral circadian clocks in adipose tissues. *Diabetes,* 55**,** 962-70.