

Explosive Expansion of $\beta\gamma$ -Crystallin Genes in the Ancestral Vertebrate

Guido Kappé · Andrew G. Purkiss ·
Siebe T. van Genesen · Christine Slingsby ·
Nicolette H. Lubsen

Received: 14 January 2010 / Accepted: 29 July 2010 / Published online: 20 August 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract In jawed vertebrates, $\beta\gamma$ -crystallins are restricted to the eye lens and thus excellent markers of lens evolution. These $\beta\gamma$ -crystallins are four Greek key motifs/two domain proteins, whereas the urochordate $\beta\gamma$ -crystallin has a single domain. To trace the origin of the vertebrate $\beta\gamma$ -crystallin genes, we searched for homologues in the genomes of a jawless vertebrate (lamprey) and of a cephalochordate (lancelet). The lamprey genome contains orthologs of the gnathostome $\beta B1$ -, $\beta A2$ - and γN -crystallin genes and a single domain γN -crystallin-like gene. It contains at least two γ -crystallin genes, but lacks the gnathostome γS -crystallin gene. The genome also encodes a non-lenticular protein containing $\beta\gamma$ -crystallin motifs, AIM1, also found in gnathostomes but not detectable in the uro- or cephalochordate genome. The four cephalochordate $\beta\gamma$ -crystallin genes found encode two-domain proteins. Unlike the vertebrate $\beta\gamma$ -crystallins but like the urochordate $\beta\gamma$ -crystallin, three of the predicted proteins contain calcium-binding sites. In the cephalochordate $\beta\gamma$ -crystallin genes, the introns are located within motif-encoding region, while in the urochordate and in the vertebrate $\beta\gamma$ -crystallin genes the introns are between motif- and/or domain encoding regions. Coincident with the evolution of the

vertebrate lens an ancestral urochordate type $\beta\gamma$ -crystallin gene rapidly expanded and diverged in the ancestral vertebrate before the cyclostomes/gnathostomes split. The β - and γN -crystallin genes were maintained in subsequent evolution, and, given the selection pressure imposed by accurate vision, must be essential for lens function. The γ -crystallin genes show lineage specific expansion and contraction, presumably in adaptation to the demands on vision resulting from (changes in) lifestyle.

Keywords $\beta\gamma$ -Crystallins · Cyclostomes · Urochordata · Cephalochordata · AIM1 · Lens evolution

Introduction

The ability to explore one's environment by vision is one of the major innovations during evolution and provides a clear selective advantage. In the vertebrate camera-type eye, it is the lens that determines visual acuity and the properties of the vertebrate lens thus correlate closely with lifestyle: aquatic and nocturnal animals have hard, round lenses, while diurnal animals have flatter and softer lenses (Land and Nilsson 2002). All vertebrate (gnathostome) eye lenses studied thus far contain the so-called ubiquitous crystallins, members of the α -, β - and γ -crystallin protein families (Bloemendal et al. 2004; Bloemendal and de Jong 1991; Wistow and Piatigorsky 1988). The properties of the lens are, amongst other factors, determined by the composition and properties of these water soluble proteins present at very high concentrations (Delaye and Tardieu 1983). At first glance, the crystallin composition of the lens appears to be rather variable. The lenses of many species contain in addition to the ubiquitous crystallins other proteins, the so-called taxon-specific crystallins, at

Electronic supplementary material The online version of this article (doi:10.1007/s00239-010-9379-2) contains supplementary material, which is available to authorized users.

G. Kappé · S. T. van Genesen · N. H. Lubsen (✉)
Biomolecular Chemistry 271, NCMLS, Radboud University
Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
e-mail: N.Lubsen@science.ru.nl

A. G. Purkiss · C. Slingsby
Department of Biological Sciences, Crystallography, Institute
of Structural and Molecular Biology, Birkbeck College, Malet
Street, London WC1E 7HX, UK

high concentrations, while lenses of other, closely related species, do not (Bloemendal et al. 2004; Piatigorsky 1989; Wistow 1993). To what extent the taxon-specific crystallins are functional in the sense of determining the exact optical properties of a lens is not clear. For example, in the gecko ι -crystallin, a taxon-specific crystallin, is likely to have been recruited as a UV filter (Werten et al. 2000). The taxon-specific crystallins are, however, merely superimposed on a common and highly conserved theme, namely the three ubiquitous crystallin protein families. The α -crystallins belong to the small heat-shock protein family (Horwitz 2003). Members of this family are found in virtually all organisms and a scenario for the evolutionary origin of the lenticular α -crystallins can easily be envisaged (de Jong et al. 1993). In contrast, the precursors of the present day vertebrate β - and γ -crystallins have been difficult to trace. The (multimeric) β - and the (monomeric) γ -crystallins are structurally and evolutionarily related: they are both built up out of four Greek key motifs organized into two domains. The basic distinction between these two gene families in gnathostomes is in the location of the introns. In the β -crystallin genes, each motif is encoded by a separate exon, while in the γ -crystallin, each domain, i.e. two motifs, is encoded by an exon (for reviews, see Bloemendal et al. 2004; Bloemendal and de Jong 1991; de Jong et al. 1993; Lubsen et al. 1988). One member of the γ -crystallin family, γ N, has an intermediate gene structure, with the N-terminal domain being encoded by a single exon (γ -crystallin gene like), while each of the two motifs of the C-terminal domain is encoded by a separate exon (β -crystallin gene like) (Weadick and Chang 2009; Wistow et al. 2005). For all β - and γ -crystallins (an additional exon(s) encodes a N-terminal extension including the start methionine).

Combining information on $\beta\gamma$ -crystallin gene organization with their encoded three-dimensional protein structures provides clues about their evolutionary origins. In the vertebrate $\beta\gamma$ -crystallins, the sequences of motifs 1 and 3 are more similar to each other than they are to motifs 2 and 4. The same holds true for motifs 2 and 4: these are more closely related to each other than they are to motifs 1 and 3. This is considered to reflect the origin of the two-domain protein from an ancient gene duplication, with present day $\beta\gamma$ -crystallin domains pairing about an approximate dyad recapitulating an ancient two-domain dimer (Blundell et al. 1981). Supporting evidence for this scenario comes from three-dimensional studies: engineered single domains can form symmetric homodimers (Norledge et al. 1996; Basak et al. 1998; Purkiss et al. 2002; Clout et al. 2000). An essential attribute of a lens protein is the ability to pack inside a lens fibre cell without forming discontinuities on the scale of half a wavelength of light. Domain pairing in β - and γ -crystallins is the first higher level of protein

organization for this superfamily. The characteristic domain pairing interaction is mediated by three key residues, donated by motifs 2 and 4. The symmetry of $\beta\gamma$ -crystallin domain pairing allows two versions of this domain assembly interaction: intramolecular pairing to form monomeric γ -crystallins, and intermolecular domain pairing by domain swapping to form certain β -crystallin dimers (Bax et al. 1990; Smith et al. 2007).

The double Greek key fold is used as the sole basic building block in the β - and γ -crystallins but also in a number of bacterial, archaeal and fungal proteins, such as in the four motif Protein S (*Mycrococcus xanthus*; Wistow et al. 1985), the two motif protein from *Methanosarcina acetivorans* (Barnwal et al. 2009) or the two motif Spherulin 3a (*Physarum polycephalum*; Kretschmar et al. 1999; see also Jaenicke and Slingsby 2001). None of these proteins are potential orthologs of the vertebrate β - and γ -crystallins. More closely related invertebrate proteins are a four motif $\beta\gamma$ -crystallin-like protein in the sponge *Geodia cydonium* (Di Maro et al. 2002) and the two motif $\beta\gamma$ -crystallin of the urochordate *Ciona intestinalis* (Shimeld et al. 2005). The gene for the *Ciona* protein is likely on the evolutionary route to the vertebrate $\beta\gamma$ -crystallin genes: it is β -crystallin-like in that the two motifs are encoded by separate exons (in contrast, the *Geodia* $\beta\gamma$ -crystallin gene lacks introns) and its promoter region drives expression to the vertebrate lens (Shimeld et al. 2005). The single domain *Ciona* protein lacks the set of hydrophobic residues involved in domain pairing on its motif 2, and it does not dimerise in solution or in the crystal lattice (Shimeld et al. 2005). The *Ciona* protein further differs from the vertebrate protein in that it has characteristic calcium-binding residues in each motif: by exploiting the approximate dyad symmetry of the domain fold, a pair of half-sites combines to form two full sites on the double motif domain. Three-dimensional studies have shown that microbial $\beta\gamma$ -crystallin-like proteins (PDB id 2k1w, Barnwal et al. 2009; PDB id 1hdf, Clout et al. 2001; PDB id 1nps, Wenk et al. 1999) have very similar calcium-binding sites as the *Ciona* protein (PDB id: 2bv2). The double Greek key fold can also be fused to other protein domains as in the epidermal differentiation-specific protein (EDSP) like proteins, thus far only detected in amphibians (Liu et al. 2008; Wistow et al. 1995), or in the Absent In Melanoma 1 (AIM1; a protein associated with suppression of malignancy of melanomas), a non-lenticular protein found in gnathostomes (jawed vertebrates). AIM1 has six $\beta\gamma$ -crystallin like domains (Ray et al. 1997) flanked at the N-terminal side by an as yet poorly defined filament-like region and at the C-terminal side by a ricin-type β -trefoil domain. In the AIM1 gene, introns are found between motif coding regions, as in the β -crystallin genes.

In gnathostomes, expression of the β - and γ -crystallin genes is mostly restricted to the lens, and tracing the

evolution of these genes would thus also shed light on the evolution of the lens. It is therefore of interest to close the gap between the two motif/one domain *Ciona* gene and the four motifs/two domain vertebrate genes. Here, we show that the gene duplications leading to the present day vertebrate β - and γ -crystallin genes preceded the divergence between the cyclostomes (jawless vertebrates) and the gnathostomes and must thus have occurred very early in vertebrate evolution. We further show that the genome of a member of the cephalochordates, *Branchiostoma floridae* (amphioxus), does encode $\beta\gamma$ -crystallin-like protein domains, which at the protein level are closely related to the *Ciona* $\beta\gamma$ -crystallin, but which have a rather different gene structure.

Materials and Methods

Searching Databases for β - and γ -Crystallin Related Sequences

The preliminary genome assembly of the *Petromyzon marinus* (http://pre.Ensembl.org/Petromyzon_marinus/index.html) and the second genome assembly of *Branchiostoma floridae* (<http://genome.jgi-psf.org/Brafl1/Brafl1.download.ftp.html>) as well as the EST databases were searched for sequences encoding β - and γ -crystallin related proteins using as query the vertebrate β - and γ -crystallin protein sequences, the *Ciona* $\beta\gamma$ -crystallin sequence as well as the sequences of EDSP and AIM1. Platypus (*Ornithorhynchus anatinus*), opossum (*Monodelphis domestica*) and armadillo (*Dasypus novemcinctus*) γ -crystallin sequences were as annotated in Ensembl and manually curated to remove some errors (for example, a Q encoded by a splice acceptor site).

Cloning of *P. marinus* β - and γ -Crystallin cDNAs

Total *P. marinus* lens RNA was reverse transcribed using the first-strand cDNA synthesis kit for RT-PCR (Roche Applied Science) according to the manufacturer's instructions with either oligo(dN)₆ primers or an oligo(dT)₁₅GC primer (5'-CCGCCCTTTTTTTTTTTTTTTT-3'). Putative β - and γ -crystallin transcripts were amplified using the primers listed in Table 1 and the Expand High Fidelity polymerase kit (Roche Applied Science). Products were ligated into the pGEM-T Easy vector (Promega) and individual inserts were sequenced using BigDye terminators and a 3730 DNA analyzer (Applied Biosystems). Pm- β A2, - β B, - γ A and - γ B are available in the Genbank database under accession numbers GQ355899, GQ355900, GQ355901 and GQ355902, respectively.

Table 1 List of primers used in the amplification of β - and γ -crystallin cDNAs from *Petromyzon marinus* eye lenses

| Name | Sequence (5'–3') | Used to amplify |
|----------|-----------------------|--------------------------------|
| 12940for | TGTGGGAGCAGGAGAACTTC | β A2-crystallin ortholog |
| 23617rev | ATGCGGCGGATGGACTGGATG | β A2-crystallin ortholog |
| 59572for | CTGGTCATCTTCGAGCTGGAG | β B1-crystallin ortholog |
| 80908rev | TTGAGAAGGTGAAGCGGCTC | β B1-crystallin ortholog |
| 1488for | ATCATCTTCTTCGAGGAGCGG | γ -crystallin |
| 1488rev | GAATCGTCGGATGCAGCGCAC | |
| 17429for | TGAATGCATCCAGGACTGTC | γ -crystallin |
| 17429rev | CGCGGATAGACTGAATCAC | |

Phylogenetic Analysis

Protein sequences were aligned using Muscle (Edgar, 2004) at default settings at the Wageningen Bioinformatics webportal (<http://www.bioinformatics.nl>). In the alignment, the AIM1 sequences were split into three four motif/two domain segments. A phylogenetic tree was inferred from the alignment using the maximum likelihood method as implemented in PhyML v3.0 (Guindon and Gascuel 2003) using the WAG substitution model, four substitution rate categories, an estimated proportion of invariable sites and gamma shape parameter. Nodal support was estimated by bootstrap analysis with 500 replicates. The sequences and accession numbers are given in the supplementary material.

Results and Discussion

β -Crystallin Related Sequences in the Lamprey

Searching the *P. marinus* (sea lamprey) genome assembly for β -crystallin related genes identified a number of contigs containing one or more β -crystallin-like exons. By sequence comparison, the predicted amino acid sequences encoded by these exons were classified as β A- or β B-crystallins (Table 2). Since none of the contigs encompassed a complete β -crystallin gene, we tried to amplify complete β -crystallin coding sequences from eye lens cDNA using primers based on the DNA sequences of the various exons encoding first and last motifs. A PCR using oligos derived from the exons on contigs 12940 and 23617 (Table 1) resulted in the amplification of a transcript representing a β A-crystallin gene (see below). In a similar fashion, the exons on contigs 59572, 50016 and 80908 could be

Table 2 List of *Petromyzon marinus* genome contigs showing $\beta\gamma$ -crystallin related sequences

| Crystallin family ^a | Contig number | Motif configuration ^b | Remarks |
|--------------------------------|---------------|--|---------|
| β A | 12940 | <M1><M2><M3> | c |
| | 23617 | <M4] | c |
| | 41226 | <M4] | |
| β B | 59572 | <M1><M2> | d |
| | 50016 | <M3> | d |
| | 80908 | <M4] | d |
| γ | 733 | <M1.M2><M3.M4] | e |
| | 1488 | <M1.M2><M3.M4] | e |
| | 11905 | <M1.M2><M3.M4] | e |
| | 17503 | <M1.M2><M3.M4] | e |
| | 24091 | <M1.M2><M3.M4] | e |
| | 41592 | <M1.M2><M3.M4] | e |
| | 1382 | <M1.M2><M3.M4] | |
| | 17429 | <M1.M2><M3.M4] | |
| | 12189 | <M1.M2><M3.M4] | f |
| γ N | 17603 | <M3><M4] | |
| | 106517 | <M1.M2> | |
| sd γ | 23770 | <M1.M2> | |
| AIM1 | 332 | <M1>.. <lt;m11>< td=""> <td>g</td> </lt;m11><> | g |

^a Classification is based on sequence comparison by phylogenetic methods and by comparing intron positions

^b Motifs are numbered, exon/intron boundaries are indicated by < and > symbols and stop-codons by *square brackets*

^c The sequence(s) found on contigs 12940 and 23617 are part of a β A-crystallin gene as validated by PCR

^d The sequence(s) found on contigs 50016, 59572 and 80908 are part of a β B-crystallin gene as validated by PCR

^e The sequence(s) found on contigs 733, 1488, 11905, 17503, 24091 and 41592 are believed to have originated by recent duplications or by assembly errors and are considered one group

^f The sequence(s) found on contig 12189 is unreliable since it contains a premature stop-codon and is flanking by a region characterized by ambiguity

^g The sequence(s) found on contig 332 contains part of the putative ortholog of the AIM1 (Absent In Melanoma 1) protein. 11 $\beta\gamma$ -motifs can be found on this contig

linked by PCR resulting in an amplified transcript representing a β B-crystallin gene. The exons encoding the N-terminal arms of the proteins could not be located with any degree of confidence in the genomic sequence and are thus not included in the amplified sequences.

γ -Crystallin Related Sequences in the Lamprey

γ -Crystallin related sequences could also be readily detected in the lamprey genome. The contigs harbouring such sequences are listed in Table 2. On most contigs two

exons were found each encoding a γ -crystallin domain and separated by an intron of approximately 1.5 kb in size. The exception is contig 17429 on which the two exons are separated by an intron of only 217 bp. As for the β -crystallin genes, the (first) exon, which encodes a 3–7 amino acid N-terminal arm in the gnathostomes could not be detected. Expression of the putative γ -crystallin genes located on contigs 1382, 1488 and 17429 in the lens was tested by PCR on eye lens cDNA. For the genes located on contigs 1488 and 17429 the corresponding transcript was found; we failed to detect a transcript from contig 1382, which could be due to differentiation and/or developmental specificity of expression. We did not test for expression of the γ -crystallin genes located on the other contigs as these are virtually identical in sequence to the γ -crystallin gene on contig 1488 and represent either very recent duplications or assembly errors.

The predicted amino acid sequences encoded by the γ -crystallin-type motifs 1 and 2 (M1.M2) exon on contig 106517 and by the two β -crystallin-type motif 3 (M3) and 4 (M4) exons linked on contig 17603 were most similar to the corresponding motifs of γ N-crystallin. We could not amplify transcripts of either the single exons or combinations thereof from eye lens cDNA, presumably because γ N-crystallin is not expressed at a high level (Wistow et al. 2005). Alignment of the predicted γ N-crystallin sequence (Pm_ γ N) with putative orthologs from zebrafish and mouse showed that the Pm_ γ N would have an insert of one amino acid in motif 3 (Fig. 1).

The *P. marinus* genome contains a second region encoding a γ N-type N-terminal domain, on contig 23770 (denoted Pm_sdG in Fig. 1). We could not detect a splice donor site either at the 3' end of the second exon region (the sequence here reads gctcg) or further downstream. The putative splice acceptor site of this potential γ N exon is directly preceded by an ATG initiation codon, suggesting that this region of the *P. marinus* genome could encode a single domain γ N-like protein lacking an N-terminal arm but with a long C-terminal tail. This putative single domain γ N-crystallin protein would have the hydrophobic domain pairing residues that allow it to form a homodimer.

Other $\beta\gamma$ -Crystallin Related Sequences in the Lamprey Genome—the AIM1 Gene

The search for β - and γ -crystallin sequences in the *P. marinus* genome yielded significant hits on contig 332. However, the predicted protein sequences aligned only poorly with fish or mammalian β - and γ -crystallins. We therefore repeated the search using AIM1 and EDSP sequences. The AIM1 sequences also matched sequences on contig 332; no potential EDSP ortholog was found. Closer inspection of contig 332 identified a total of 13 exons

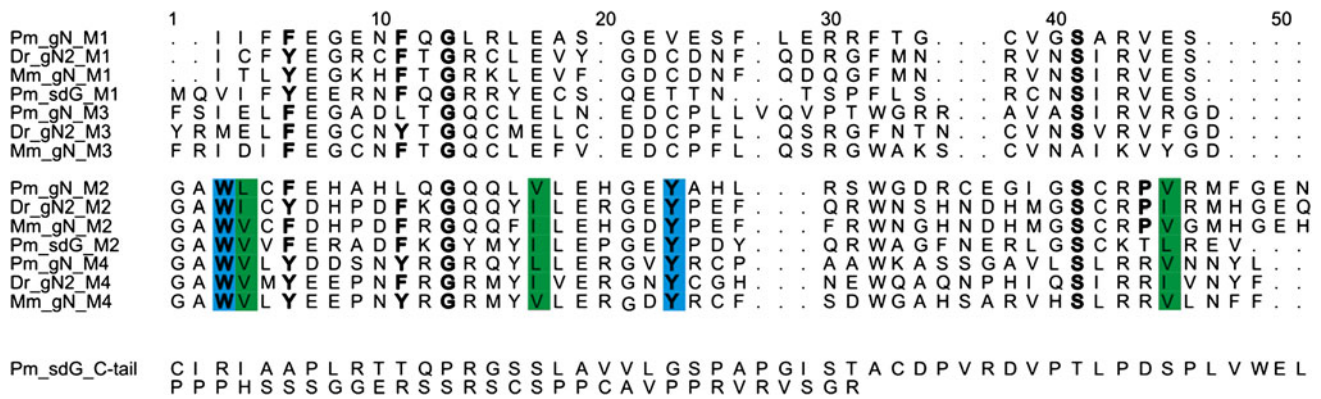


Fig. 1 The *P. marinus* γ N protein. The predicted sequences of Pm γ N-crystallin and of the putative γ N-crystallin-type single domain protein (Pm_sdG) are aligned with the *D. rerio* and *M. musculus* γ N-crystallin sequence. Structurally important residues are in bold, the conserved tyrosine and tryptophan corners in blue, and the three

hydrophobic residues in motifs 2 and 4 involved in domain pairing in green. The predicted sequence of the C-terminal extension of Pm_sdG is given below the alignment. The alignment is split in the odd and even motifs

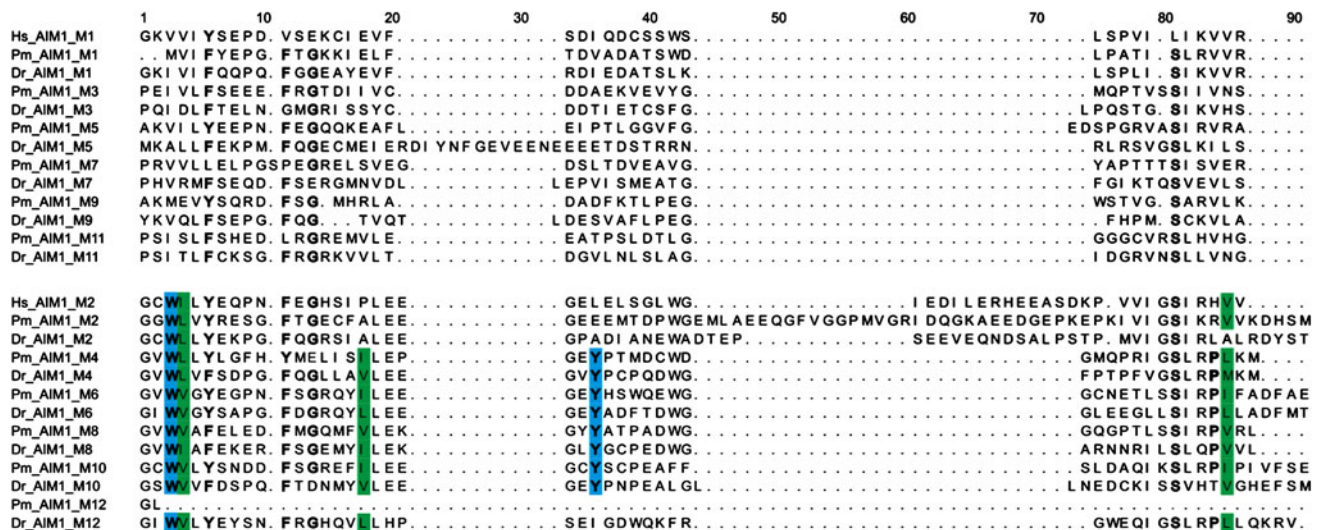


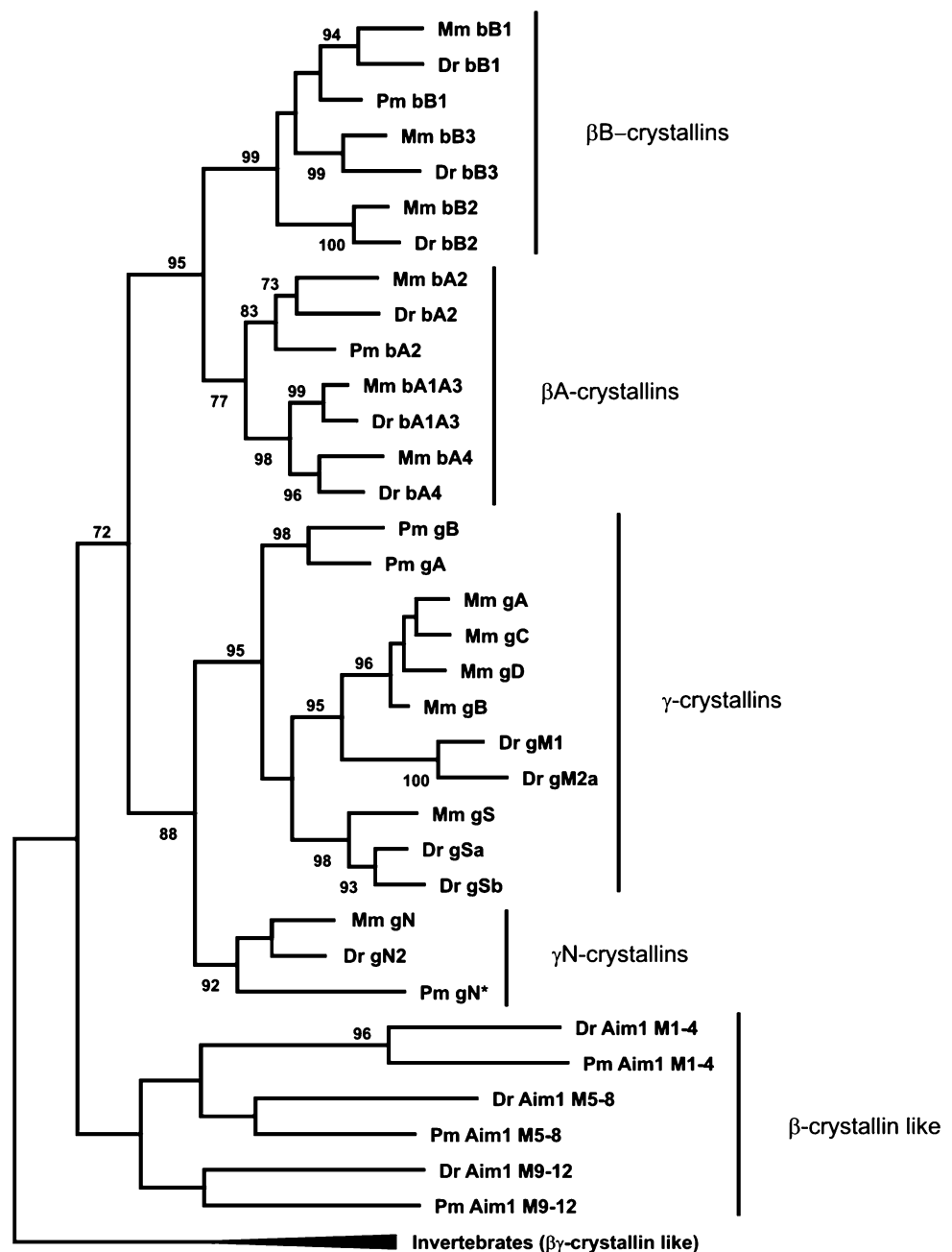
Fig. 2 Alignment of *B. floridae* $\beta\gamma$ -crystallin related sequences. The predicted protein sequences of *B. floridae* Bf-bg1, -bg2, -bg3 and -bg4 (acc.nr. FE565747, BW704196 and BW723025) are aligned with the *C. intestinalis* $\beta\gamma$ -crystallin (Ci_bg) and the *D. rerio* β B2- (Dr_bB2) and γ Sa-crystallin (Dr_gSa). The motifs are indicated as M followed by the number. Structurally important residues are indicated as in

Fig. 1; calcium-binding residues are in grey. The sequence in italics in Bf-bg3 M1 was derived from the genomic sequence; the C-terminal extensions of Bf-bg1 and Bf-bg2 are shown below the alignment. The # indicates the position of the HVNPANT insert in Bf-bg1 M2 and & that of PPSNM in Bf-bg1 M4. The alignment is split in the odd and even motifs

together encoding AIM1-like $\beta\gamma$ -crystallin motifs 1-11 and two linker regions (Fig. 2). The exons encoding the twelfth $\beta\gamma$ -crystallin motif, the C-terminal ricin domain and the N-terminal half of the AIM1 protein could not be found. The intron positions and phases of the *P. marinus* AIM1 exons are identical to the ones in the human AIM1 gene. An alignment of the 11 Pm AIM1 $\beta\gamma$ -crystallin motifs with the corresponding motifs from the *D. rerio* and, for comparison, motifs 1 and 2 from human AIM1 clearly shows Pm AIM1 is the ortholog of the AIM1 gene in gnathostomes (Fig. 2, see also Fig. 3). The large insert in the second Pm $\beta\gamma$ -crystallin motif between positions 40 and 70 is present in

all second AIM1 $\beta\gamma$ -crystallin motifs. Determination of the structure of the first domain of human AIM1 protein shows that the tertiary structure stays intact despite the extra bulge within the second motif (Aravind et al. 2008). The seventh Pm $\beta\gamma$ -crystallin motif does deviate substantially from the others since it lacks most of the conserved residues that specify the folded hairpin of the Greek key motif-fold (shown in bold black typeface in Fig. 2). The presence of an AIM1 gene in *P. marinus* shows that the origin of this gene must predate the cyclostomes–gnathostomes split. We could not find an AIM1 gene in *B. floridae* or in *C. intestinalis* (Shimeld et al. 2005; unpubl. res. and see below).

Fig. 3 Evolution of the vertebrate $\beta\gamma$ -crystallins. A maximum likelihood tree based on the alignment of both vertebrate and invertebrate $\beta\gamma$ -crystallin sequences (see supplementary material) was constructed. Species are indicated as Mm *M. musculus*, Dr *D. rerio* and Pm *P. marinus*. Invertebrate $\beta\gamma$ -crystallins were included in building the tree; for clarity sake the invertebrate branches have been collapsed (arrow). The Pm_gN is predicted from genomic sequence and marked with an asterisk, other Pm $\beta\gamma$ -crystallin sequences are as validated by RT-PCR. The Dr and predicted Pm AIM1 sequences were divided into three four motif, two-domain segments as indicated. Numbers adjacent to nodes indicate percentage bootstrap support; only values larger than 70% (of 500) are shown



This suggests emergence of the AIM1 gene after the urochordate and vertebrate divergence.

Phylogenetic Analysis of the *P. marinus* $\beta\gamma$ -Crystallins and Related Proteins

On the basis of an alignment of the deduced sequences of the *P. marinus* β - and γ -crystallins with representative gnathostome β - and γ -crystallins sequences as well as AIM1 motifs and invertebrate $\beta\gamma$ -crystallins (see below and supplementary material) the phylogenetic tree shown in Fig. 3 was inferred. The *P. marinus* AIM1 $\beta\gamma$ -crystallin domains

and the vertebrate (lens) $\beta\gamma$ -crystallins form separate clades. This strongly suggests that these proteins originated independently from an ancestral $\beta\gamma$ -crystallin sequence. Within the $\beta\gamma$ -crystallin branch, the two Pm β -crystallin sequences cluster with β A2-crystallin and with β B1-crystallin, respectively. The tree shows that the gene duplications leading first of all to the acidic and basic β -crystallins and subsequently to the subdivision in β A2- and β A3/ β A4-crystallin within the acidic β -crystallins and β B1- and β B2/ β B3-crystallins within the basic β -crystallins must have preceded the divergence between cyclostomes and the gnathostomes. One expects then to find genes for a

β A3/ β A4-crystallin and a β B2/ β B3-crystallin in the *P. marinus* genome as well. The genome sequence is about 80% complete (pre.ensembl.org/Petromyzon_marinus/Info/StatsTable) and those genes could thus still be missing from the sequence. Furthermore, the lamprey genome is extensively rearranged in somatic cells during development and genes could thus be present in the germ line genome but lacking from the DNA of somatic cells (Smith et al. 2009) and from the present genome assembly. The orphan β A-crystallin-type motif 4 on contig 41226 (Table 2) does suggest that genomic information is still missing. The alternative is that the genes have been lost again during the evolution of the cyclostomes.

The phylogenetic tree also shows that the putative Pm γ N-crystallin groups with the mouse and fish γ N-crystallin, whereas the Pm γ A- and γ B-crystallin genes form a separate clade. The *P. marinus* genome appears to lack the γ S-crystallin gene common to gnathostomes. Either this gene originated after divergence of the cyclostomes and the gnathostomes or it is located in that part of the *P. marinus* genome that still needs to be sequenced and assembled. It has been previously shown that the γ S-crystallin genes in the gnathostomes are orthologs. Similarly, the γ N-crystallin genes all descended from the same ancestral gene (see also Fig. 4). In contrast, the other γ -crystallin genes have repeatedly expanded and contracted (Lubsen et al. 1988; Weadick and Chang 2009; Wistow et al. 2005). To determine when the mammalian crygA-F genes originated, we added platypus (*O. anatinus*), opossum (*M. domestica*) and armadillo (*D. novemcinctus*), *Xenopus laevis*, rat (*Rattus norvegicus*) and human (*Homo sapiens*) γ -crystallin sequences to the phylogenetic tree. As shown in Fig. 4, the mammalian γ A-F radiation from a single γ -crystallin gene preceded the earliest split in the mammalian lineage, that between Prototheria and Theria. The selective pressure on the variable subclass of the γ -crystallins is thought to result from the adaptation of the lens shape and properties to the lifestyle. A high γ -crystallin level correlates with low lens water content (for review, see Bloemendal et al. 2004) and thus round, hard lenses suitable for aquatic or nocturnal life. The soft chicken lens, for example, contains only γ N- and γ S-crystallin and lacks the variable subgroup of γ -crystallins altogether (but does have large amounts of a taxon-specific enzyme crystallin). Fish eyes have steep, symmetric gradients of refractive index, approximately parabolic in shape, to increase focusing power whilst correcting for spherical aberration (Land and Nilsson 2002). They are often multifocal, which corrects for chromatic aberration, allowing high-resolution colour vision (Kröger et al. 1999). The positive and negative corrections for spherical aberration, which facilitate spectral tuning to the retinal photopigments, are considered to stem from small perturbations to the shape of the lens refractive index

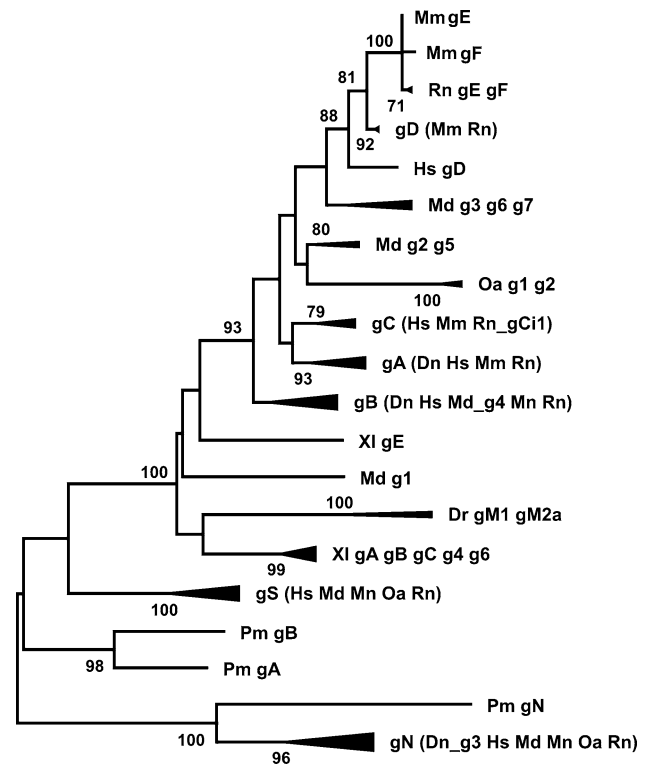


Fig. 4 Evolution of the γ -crystallins. A maximum likelihood tree based on an alignment of γ -crystallins (supplementary material) was constructed. Species are indicated as in the legend to Fig. 3. In addition, sequences were used from platypus (*Ornithorhynchus anatinus* Oa), opossum (*Monodelphis domestica* Md), armadillo (*Dasyurus novemcinctus* Dn), rat (*Rattus norvegicus* Rn), *Xenopus laevis* (Xl) and man (*Homo sapiens* Hs). Sequences and accession numbers can be found in the supplementary material. For clarity sake, some branches have been collapsed (arrows). gA, gB and gC denote the corresponding genes from Hs, Mm, Rn, and Dn, gS the genes from Hs, Mm, Rn, Md and Oa, gN the genes from Hs, Mm, Rn, Md, Dn and Oa. Numbers adjacent to nodes indicate percentage bootstrap support; only numbers larger than 70 are shown

gradient. Multifocality may be of ancient evolutionary origin in vertebrates (Karpestam et al. 2007), and recently it has been shown that in the adult stage, the *P. marinus* has a multifocal lens (Gustafsson et al. 2008). High levels of sulphur containing residues may contribute towards a high power lens by increasing the protein refractive index increment and/or the ability to close pack. The fish γ M proteins contain high levels of sulphur containing residues, particularly methionine (Chang et al. 1988): for example for *D. rerio*, out of the (177) amino acid residues of γ M1 (excluding the start methionine), 15.2% are cysteine or methionine, of γ M2a 18.9%, of γ M3 13.3%, of γ M4 11.0%, of γ M5 13.0%, of γ M6 10.2% and of γ M7 13.8%; however, in human lens the sulphur level for γ C is 7.5% and for γ D is 5.8%. By comparison, in the lamprey 172-residue γ A sequence, the sulphur level is 9.9%, and for γ B it is 8.8%, while of the two elasmobranch lipshark γ M sequences, M1 has 21.2% sulphur and M2 has 9.7%. The lens focal length

(normalized for lens radius) of the deep water *P. marinus* (2.31R) is in the lower range for teleost lenses, which is 2.2–2.8R (Gustafsson et al. 2008). It would be useful to collect clade specific γ -crystallin sequences, measure their refractive index increments, and compare these with the measured lens optical properties.

$\beta\gamma$ -Crystallin Related Sequences in Invertebrates

We have previously identified a single domain $\beta\gamma$ -crystallin in the urochordate *C. intestinalis* (Shimeld et al. 2005). To determine whether cephalochordates also contain a similar $\beta\gamma$ -crystallin gene, we searched the genome assembly of the cephalochordate *Branchiostoma floridae* (amphioxus) for sequences encoding $\beta\gamma$ -crystallin related motifs. This search did not yield potential orthologs of AIM1 and EDSP but did yield four $\beta\gamma$ -crystallin related genes. Three of these, Bf-bg1, Bf-bg2 and Bf-bg3 (Fig. 5), are each supported by a single EST. Bf-bg2 and Bf-bg3 are closely

linked head to tail. Gene models predict an alternative transcript of this genome region, which would contain the first three exons of Bf-bg2 and the first exon present in the Bf-bg3 EST (see Fig. 6; note that the Bf-bg3 EST is likely to be incomplete at the 5' end, see below). This fourth possible $\beta\gamma$ -crystallin coding sequence is here denoted as Bf-bg4. The EST coverage of this region is sparse and insufficient to exclude the possibility of Bf-bg4. Furthermore, there may still be problems with the genome assembly as evidenced by the fact that the Bf-bg3 region is inverted in the second assembly of the genome relative to the first assembly. The Bf-bg1, Bf-bg2 and Bf-bg4 protein sequences have four $\beta\gamma$ -crystallin motifs encoded by four exons. However, unlike in the *Ciona*, or vertebrate β - and γ -crystallin genes and the $\beta\gamma$ -crystallin-like region in the AIM1 gene, introns 1 and 2 are located within and not between the motif coding regions (see below). If the AUG codon in exon one is indeed the initiation codon, then these proteins would also lack the short N-terminal arm encoded

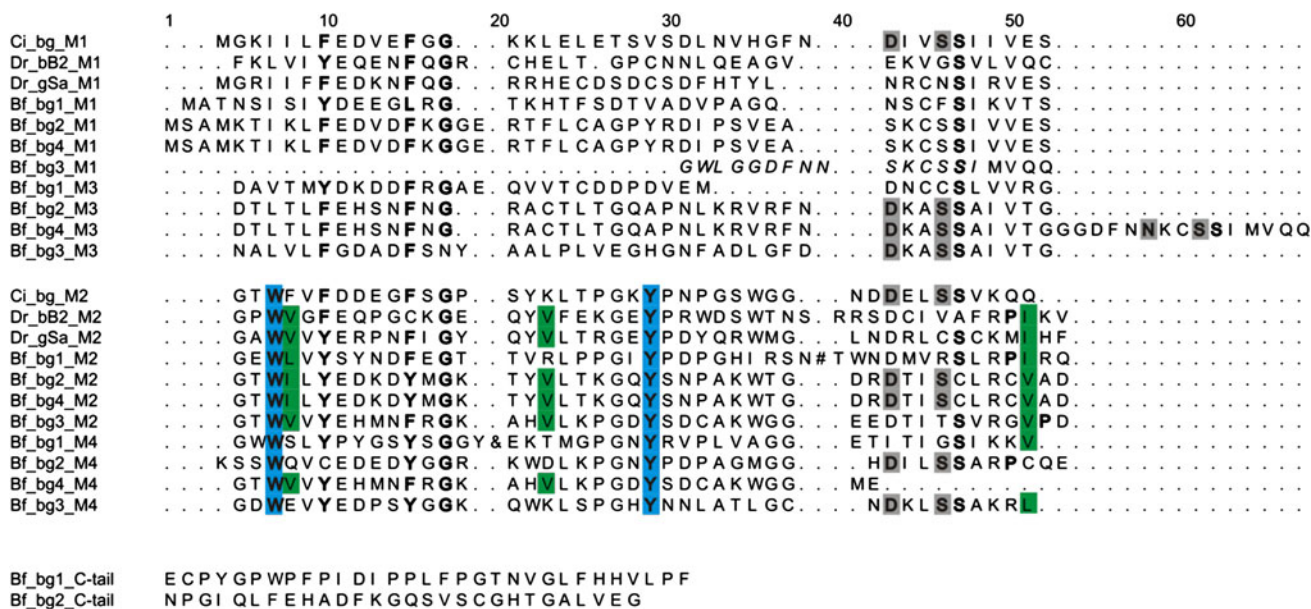


Fig. 5 Alignment of *P. marinus* AIM1 $\beta\gamma$ -crystallin motifs. The AIM1 $\beta\gamma$ -crystallin motifs found in the *P. marinus* (Pm) genome assembly are aligned with those of the *D. rerio* (Dr) AIM1 protein and

with the first two $\beta\gamma$ -crystallin motifs of human (Hs) AIM1. The alignment is split in the odd and even motifs. Structurally important residues are indicated as in Fig. 1

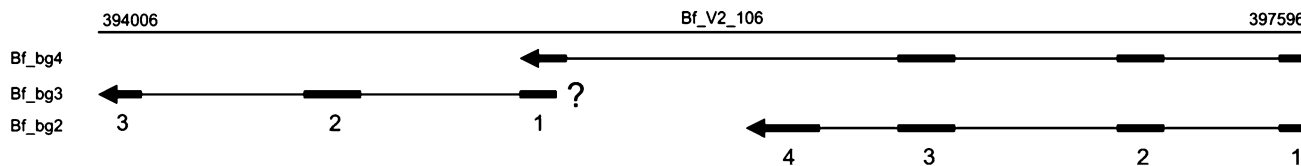


Fig. 6 Schematic representation of the genomic arrangement of the exons of Bf-bg2, Bf-bg3 and Bf-bg4. The ESTs mapped to nucleotide region 394006–397596 of contig 106 (version 2 of the genomic assembly) are shown with the exons indicated by the *thick lines* and

the introns by the *thin lines*. The *arrowheads* show the direction of transcription, the numbers indicate the exon numbering as referred to in the text and the ? indicates that this exon is likely incomplete at the 5' end

by a separate 5' exon in all the vertebrate β - and γ -crystallins genes as well as in the *Ciona* $\beta\gamma$ -crystallin gene. The Bf-bg3 sequence, as far as it can be deduced from the corresponding EST, is very similar to that of Bf-bg2, but would be missing the first motif and is probably incomplete.

An alignment of the Bf-bg protein sequences shows the same pattern of motif similarity as vertebrate $\beta\gamma$ -crystallins: motifs 2 and 4 are more similar to each other than to motifs 1 and 3 (Fig. 5). However, the divergence between Bf-bg2 or Bf-bg3 motifs is less than in the vertebrate β - and γ -crystallins, suggesting a recent duplication from a one domain to a two domain encoding gene. The critical Greek key motif residues are also conserved in the Bf-bg proteins (Fig. 5, in bold) with few exceptions, such as the highly conserved G replaced by an N in motif 3 of Bf-bg3, and the highly conserved S replaced by C in Bf-bg2 motif 2 (Fig. 5). Lens $\beta\gamma$ -crystallin domains have a conserved tyrosine and tryptophan corner, both residues being contributed by motifs 2 and 4 for each domain. These residues are conserved in equivalent motifs in *Ciona* protein and the Bf-bg proteins (Fig. 5, in blue). The calcium-binding residues are also conserved in both motifs 3 and 4 of Bf-bg2 and 3 (indicated in grey in Fig. 5), making it likely that the C-terminal domains in these proteins bind calcium, as does the single domain *Ciona* $\beta\gamma$ -crystallin. The Bf-bg4 sequence lacks the motif 4 part of the calcium-binding site and would thus be predicted not to bind calcium. The second and fourth motifs of Bf-bg1 have an insert; that in the second is at the same position as the insert in the second motif of the first $\beta\gamma$ -crystallin domain of AIM1 (Fig. 2).

The two domain Bf-bg proteins lack the domain pairing hydrophobics in motif 4. Motifs 2 of Bf-bg2 and Bf-bg3 each do have the full set, leading to the possibility that these N-terminal domains might pair with each other (to form homo- or heterodimers), leaving their C-terminal domains unpaired.

A protein with a sequence distantly related to $\beta\gamma$ -crystallins has been characterized from the phylum Porifera, the sponge *Geodia cydonium* (Krasko et al. 1997; Giancola et al. 2005). The completed genome of the cnidarian sea anemone, *Nematostella vectensis*, encodes a distant $\beta\gamma$ -crystallin relative that groups in the phylogenetic tree with the sponge sequence, as does another cnidarian hexacorrallia sequence, reported from the coral *Montipora capitata* (Fig. 7). The cnidarian sequences attributed to motifs 3 and 4 each have the characteristic amino acid residues for the Greek key hairpin fold, with *N. vectensis* having both the tyrosine and tryptophan corner residues conserved in motif 4, thus providing good evidence that the cnidarian C-terminal domains will have the double Greek key $\beta\gamma$ -crystallin fold. The sponge motif 3 sequence is short and in the absence of three-dimensional information, it is unclear how this motif completes the double Greek key fold. The motif 2 sequences

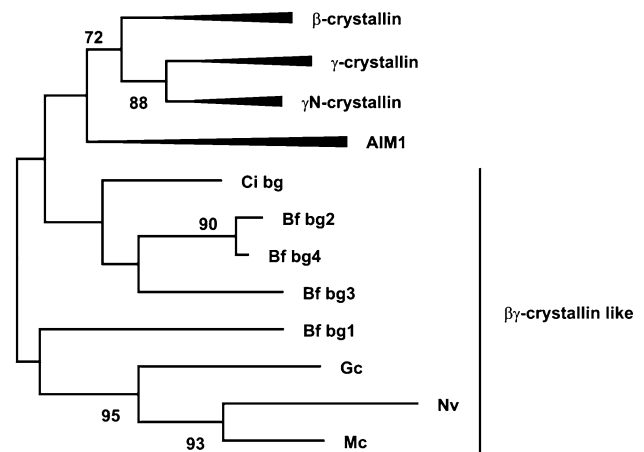


Fig. 7 Evolution of the $\beta\gamma$ -crystallins. A maximum likelihood tree based on the alignment of both vertebrate and invertebrate $\beta\gamma$ -crystallin sequences (see supplementary material) was constructed. Species are indicated as Ci *C. intestinalis*, Bf *B. floridae*, Gc *Geodia cydonium*, Mc *Montipora capitata* and Nv *Nematostella vectensis*. For clarity sake, the vertebrate branches of the tree have been collapsed; this part of the tree is shown in Fig. 3. Numbers adjacent to nodes indicate percentage bootstrap support; only values larger than 70% (of 500) are shown

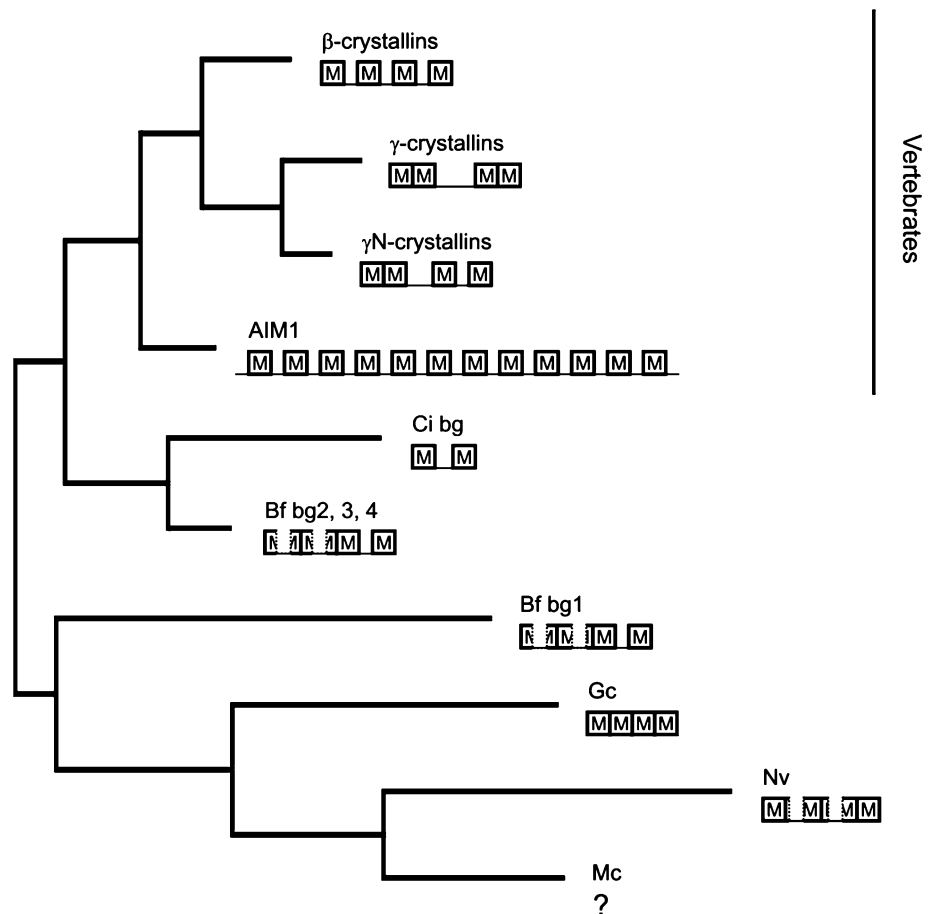
(from the three N-terminal domains) have most of the conserved residues, whereas the motif 1 sequences have hardly any, making an alignment unreliable. It is interesting that the camera eye of the cnidarian jelly fish recruited enzymes for a lens role (Piatigorsky and Kozmik 2004) even though $\beta\gamma$ -crystallin-like proteins were present in the phylum.

Origin of the Vertebrate $\beta\gamma$ -Crystallin Gene Family

The phylogenetic trees presented in Figs. 3 and 7 clearly show that the present day vertebrate β - and γ -crystallin gene family must already have been present in the last common ancestor of the cyclostomes and the gnathostomes. If we superimpose the gene structure on the phylogenetic tree (Fig. 8) the obvious hypothesis is that the ancestral vertebrate $\beta\gamma$ -crystallin gene was a single domain *C. intestinalis* like gene, with each motif being encoded by a single exon. Overall, the comparison of the $\beta\gamma$ -crystallin coding sequence and gene structures is in line with the tree calculated from 146 nuclear genes showing that urochordates are closer to vertebrates than cephalochordates (Delsuc et al. 2006).

The common theme in $\beta\gamma$ -crystallins is a four motif/two domain structure. Protein structure and sequence similarity between domains strongly suggest that the gene encoding the two domain protein originated from two successive duplications: the duplication of a motif coding segment to a single domain encoding gene and duplication of the single domain gene to a two domain gene (Fig. 8). The only known eukaryotic representatives of the presumptive ancestral

Fig. 8 Changes in structure of the $\beta\gamma$ -crystallin genes during evolution. Exonic regions encoding a motif are indicated as a boxed *M*, introns are shown by a line. Only the exonic regions encoding the domains are shown. The gene structure of the *Mc* $\beta\gamma$ -crystallin gene is unknown; the sequence reported is an EST



single domain gene are the *C. intestinalis* $\beta\gamma$ -crystallin and *P. polycephalum* Spherulin 3a genes. Our hypothesis is that gene expansion from a *Ci*-type $\beta\gamma$ -crystallin single domain gene accompanied by fusion and shifts in intron position has occurred repeatedly in the various eukaryotic lineages. Even though intron positions tend to be conserved, lineage specific loss and gain is known to have happened during eukaryotic evolution (Carmel et al. 2007; Rogozin et al. 2003; Scott and Gilbert 2006).

A *Ci* $\beta\gamma$ -crystallin-like gene is not only the likely ancestor of the vertebrate $\beta\gamma$ -crystallin genes but also of the AIM1 gene: the intron positions correspond exactly. In the ancestral vertebrate, the ancestral $\beta\gamma$ -crystallin gene underwent an explosive series of gene duplications and gene fusions to yield an AIM1 gene, an ancestral β -crystallin gene, an ancestral γ -crystallin gene and a γ N-crystallin gene. It has been suggested that the γ -crystallin gene is a descendant of a β -crystallin gene, which underwent successive loss of the between-motif introns. The γ N-crystallin gene, with a γ -crystallin-type exon encoding the N-terminal domain and two β -crystallin-like exons encoding the C-terminal domain would be the retained intermediate (Wistow et al. 2005). Our results do neither support nor negate this hypothesis: both types of genes were already

present in the ancestral vertebrate. Our results also do not support or negate the alternative hypothesis, namely that the intron between motifs was first lost from a *Ciona* type single domain encoding gene, which then duplicated and fused to form a γ -crystallin gene. The single domain intron-less γ -like gene in the *P. marinus* genome could represent a remnant of a putative single domain γ -crystallin precursor gene; it could equally well be the result of a mutation in the splice donor site of a γ N-crystallin gene. The β -crystallin gene in the ancestral vertebrate must have duplicated and diversified further to the family of β -crystallin genes found in all vertebrate genomes. In contrast, the phylogenetic tree (Figs. 3, 4) suggests that the γ -crystallin gene remained single and duplicated only after divergence of the various vertebrate lineages. If so, the ancestral vertebrate lens is likely to have had a low level of γ -crystallin, which, extrapolating backwards from the properties of the present day vertebrate lenses, would indicate that the lens had a high water content, was soft and only a bit higher in refractive index than non-lens cells.

$\beta\gamma$ -Crystallin-like genes encoding double Greek key folds are found in several microbial forms including bacteria, archaea and a eumycetozoa, but have not yet been found in several of the major animal phyla such as

nematoda, arthropoda, mollusca, platyhelminthes, enteropneusta and echinodermata. It may be that despite the three-dimensional structural similarities between microbial and vertebrate $\beta\gamma$ -crystallins, the genes did arise independently in the various lineages. Alternatively, loss of these genes may have been a not uncommon event in prevertebrate evolution. Although vision is common in animals, what is innovatory in vertebrates is a camera-type eye with the retina derived from ciliary photoreceptors, something we share with jellyfish, but they lack the brain (Lamb et al. 2007). The genome sequences show that there has been massive expansion of the $\beta\gamma$ -crystallin family in vertebrates, with exon encoding of the motifs providing the necessary flexibility. The driving force behind this expansion would be the remarkable adaptations in the optical systems for different lifestyles, even amongst different teleosts that eat different kinds of food, at different depths, under different light levels (Karpestam et al. 2007). The clade-specific γ -crystallins contribute at some level to the required variations in refractive index gradients; the ability to vary the levels of specific proteins, something in which lenses appear to excel, is useful as well.

Acknowledgments CS gratefully acknowledges the financial support of the Medical Research Council, London (Grant Number G0801846).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aravind P, Wistow G, Sharma Y, Sankaranarayanan R (2008) Exploring the limits of sequence and structure in a variant $\beta\gamma$ -crystallin domain of the protein Absent In Melanoma-1 (AIM1). *J Mol Biol* 381:509–518
- Barnwal RP, Jobby MK, Devi KM, Sharma Y, Chary KVR (2009) Solution structure and calcium-binding properties of m-crystallin, a primordial $\beta\gamma$ -crystallin from archaea. *J Mol Biol* 386:675–689
- Basak AK, Kroone RC, Lubsen NH, Naylor CE, Jaenicke R, Slingsby C (1998) The C-terminal domains of gamma S-crystallin pair about a distorted twofold axis. *Prot Eng* 11:337–344
- Bax B, Lapatto R, Nalini V, Driessen H, Lindley PF, Mahadevan D, Blundell TL, Slingsby C (1990) X-ray analysis of beta B2-crystallin and evolution of oligomeric lens proteins. *Nature* 347:776–780
- Bloemendal H, de Jong WW (1991) Lens proteins and their genes. *Prog Nucl Acid Res Mol Biol* 41:259–281
- Bloemendal H, de Jong W, Jaenicke R, Lubsen NH, Slingsby C, Tardieu A (2004) Ageing and vision: structure, stability and function of lens crystallins. *Prog Biophys Mol Biol* 86:407–485
- Blundell T, Lindley P, Miller L, Moss D, Slingsby C, Tickle I, Turnell B, Wistow GJ (1981) The molecular structure and stability of the eye lens: X-ray analysis of gamma-crystallin II. *Nature* 289:771–777
- Carmel L, Rogozin IB, Wolf YI, Koonin EV (2007) Patterns of intron gain and conservation in eukaryotic genes. *BMC Evol Biol* 7:192
- Chang T, Jiang Y-J, Chiou S-H, Chang W-C (1988) Carp gamma-crystallins with high methionine content: cloning and sequencing of the complementary DNA. *Biochim Biophys Acta* 951:226–229
- Clout NJ, Basak A, Wieligmann K, Bateman OA, Jaenicke R, Slingsby C (2000) The N-terminal domain of [beta]b2-crystallin resembles the putative ancestral homodimer. *J Mol Biol* 304:253–257
- Clout NJ, Kretschmar M, Jaenicke R, Slingsby C (2001) Crystal structure of the calcium-loaded Spherulin 3a dimer sheds light on the evolution of the eye lens betagamma-crystallin domain fold. *Structure* 9:115–124
- de Jong W, Leunissen J, Voorter C (1993) Evolution of the alpha-crystallin/small heat-shock protein family. *Mol Biol Evol* 10:103–126
- Delage M, Tardieu A (1983) Short-range order of crystallin proteins accounts for eye lens transparency. *Nature* 302:415–417
- Delsuc F, Brinkmann H, Chourrout D, Philippe H (2006) Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439:965–968
- Di Maro A, Pizzo E, Cubellis MV, D'Alessio G (2002) An intron-less [beta][gamma]-crystallin-type gene from the sponge *Geodia cydonium*. *Gene* 299:79–82
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Giancola C, Pizzo E, Di Maro A, Cubellis MV, D'Alessio G (2005) Preparation and characterization of geodin—a beta gamma-crystallin-type protein from a sponge. *FEBS J* 272:1023–1035
- Guindon S, Gascuel OA (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704
- Gustafsson OSE, Collin SP, Kröger RHH (2008) Early evolution of multifocal optics for well-focused colour vision in vertebrates. *J Exp Biol* 211:1559–1564
- Horwitz J (2003) Alpha-crystallin. *Exp Eye Res* 76:145–153
- Jaenicke R, Slingsby C (2001) Lens crystallins and their microbial homologs: structure, stability, and function. *Crit Rev Biochem Mol Biol* 36:435–499
- Karpestam B, Gustafsson J, Shashar N, Katzir G, Kröger RHH (2007) Multifocal lenses in coral reef fishes. *J Exp Biol* 210:2923–2931
- Krasko A, Muller IM, Muller WEG (1997) Evolutionary relationships of the metazoan beta gamma-crystallins, including that from the marine sponge *Geodia cydonium*. *Proc R Soc Lond B* 264:1077–1084
- Kretschmar M, Mayr E, Jaenicke R (1999) Homo-dimeric spherulin 3a: a single-domain member of the beta gamma-crystallin superfamily. *Biol Chem* 380:89–94
- Kröger RHH, Campbell MCW, Fernald RD, Wagner H-J (1999) Multifocal lenses compensate for chromatic defocus in vertebrate eyes. *J Comp Phys* 184:361–369
- Lamb TD, Collin SP, Pugh EN (2007) Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. *Nat Rev Neurosci* 8:960–976
- Land MF, Nilsson D-E (2002) *Animal eyes*. Oxford University Press, Oxford
- Liu S-B, He Y-Y, Zhang Y, Lee W-H, Qian J-Q, Lai R, Jin Y (2008) A novel non-lens betagamma-crystallin and trefoil factor complex from amphibian skin and its functional implications. *PLoS ONE* 3:e1770
- Lubsen NH, Aarts HJM, Schoenmakers JGG (1988) The evolution of lenticular proteins—the beta-crystallin and gamma-crystallin super gene family. *Prog Biophys Mol Biol* 51:47–76
- Norledge BV, Mayr EM, Glockshuber R, Bateman OA, Slingsby C, Jaenicke R, Driessen HP (1996) The X-ray structures of two

- mutant crystallin domains shed light on the evolution of multi-domain proteins. *Nat Struct Biol* 3:267–274
- Piatigorsky J (1989) Lens crystallins and their genes: diversity and tissue-specific expression. *FASEB J* 3:1933–1940
- Piatigorsky J, Kozmik Z (2004) Cubozoan jellyfish: an Evo/Devo model for eyes and other sensory systems. *Int J Dev Biol* 48:719–729
- Purkiss AG, Bateman OA, Goodfellow JM, Lubsen NH, Slingsby C (2002) The X-ray crystal structure of human gamma S-crystallin C-terminal domain. *J Biol Chem* 277:4199–4205
- Ray ME, Wistow G, Su YA, Meltzer PS, Trent JM (1997) AIM1, a novel non-lens member of the $\beta\gamma$ -crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. *Proc Natl Acad Sci USA* 94:3229–3234
- Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV (2003) Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. *Curr Biol* 13:1512–1517
- Scott WR, Gilbert W (2006) The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat Rev Genet* 7:211–221
- Shimeld SM, Purkiss AG, Dirks RPH, Bateman OA, Slingsby C, Lubsen NH (2005) Urochordate $\beta\gamma$ -crystallin and the evolutionary origin of the vertebrate eye lens. *Curr Biol* 15:1684–1689
- Smith MA, Bateman OA, Jaenicke R, Slingsby C (2007) Mutation of interfaces in domain-swapped human betaB 2-crystallin. *Protein Sci* 16:615–625
- Smith JJ, Antonacci F, Eichler EE, Amemiya CT (2009) Programmed loss of millions of base pairs from a vertebrate genome. *Proc Natl Acad Sci USA* 106:11212–11217
- Weadick CJ, Chang BSW (2009) Molecular evolution of the $\beta\gamma$ lens crystallin superfamily: evidence for a retained ancestral function in γ N crystallins? *Mol Biol Evol* 26:1127–1142
- Wenk M, Baumgartner R, Holak TA, Huber R, Jaenicke R, Mayr E-M (1999) The domains of protein S from *Myxococcus xanthus*: structure, stability and interactions. *J Mol Biol* 286:1533–1545
- Werten PJJ, Röhl B, van Aalten DMF, de Jong WW (2000) Gecko ι -crystallin: how cellular retinol-binding protein became an eye lens ultraviolet filter. *Proc Natl Acad Sci USA* 97:3282–3287
- Wistow GJ (1993) Lens crystallins: gene recruitment and evolutionary dynamism. *Trends Biochem Sci* 18:301–306
- Wistow GJ, Piatigorsky J (1988) Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. *Ann Rev Biochem* 57:479–504
- Wistow G, Summers L, Blundell T (1985) *Myxococcus xanthus* spore coat protein S may have a similar structure to vertebrate lens beta gamma-crystallins. *Nature* 315:771–773
- Wistow G, Jaworski C, Rao PV (1995) A non-lens member of the beta gamma-crystallin superfamily in a vertebrate, the amphibian *Cynops*. *Exp Eye Res* 61:637–639
- Wistow G, Wyatt K, David L, Gao C, Bateman O, Bernstein S, Tomarev S, Segovia L, Slingsby C, Vihtelic T (2005) gammaN-crystallin and the evolution of the betagamma-crystallin superfamily in vertebrates. *FEBS J* 272:2276–2291