
Shimeld, Sebastian M.; Purkiss, Andrew G.; Slingsby, Christine et al (2005). Urochordate beta gamma-crystallin and the evolutionary origin of the vertebrate eye lens. *Current Biology* 15 (18) 1684-1689.

Urochordate βγ-crystallin and the evolutionary origin of the vertebrate eye lens.

Sebastian M. Shimeld,[1] Andrew G. Purkiss,[2] Ron P.H. Dirks,[3] Orval A. Bateman,[2] Christine Slingsby[2] and Nicolette H. Lubsen[3].

1. Department of Zoology, University of Oxford, The Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK

2. Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

3. Department of Biochemistry, Faculty of Science, Radboud University Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

Corresponding author: Sebastian Shimeld

Email: Sebastian.Shimeld@zoo.ox.ac.uk

Running Head: *Ciona βγ-crystallin* and lens evolution

Summary

A refracting lens is a key component of our image-forming camera eye, however its evolutionary origin is unknown, as precursor structures appear absent in non-vertebrates [1]. The vertebrate βγ-crystallin genes encode abundant structural proteins critical for the function of the lens [2]. We show that the urochordate *Ciona intestinalis*, which split from the vertebrate lineage before the evolution of the lens, has a single gene coding for a single domain monomeric βγ-crystallin. The crystal structure of *Ciona*-βγ-crystallin is very similar to that of a vertebrate βγ-crystallin domain, except for paired, occupied calcium binding sites. The *Ciona*-βγ-crystallin is only expressed in the palps and in the otolith, the pigmented sister cell of the light sensing ocellus. The *Ciona-βγ crystallin* promoter region targeted expression to the visual system, including lens, in transgenic *Xenopus* tadpoles. We conclude that the vertebrate βγ-crystallins evolved from a single domain protein already expressed in the neuroectoderm of the pre-vertebrate ancestor. The conservation of the regulatory hierarchy controlling βγ-crystallin expression between organisms with and without a lens shows that the evolutionary origin of the lens was based on co-option of pre-existing regulatory circuits controlling the expression of a key structural gene in a primitive light sensing system.

Most living vertebrates possess anterior paired eyes, each with a lens. While anterior photoreceptors are known to have evolved before the radiation of the major lineages of bilaterally symmetrical animals [3, 4], the vertebrate lens is a more recent innovation that evolved in the vertebrate lineage. Indeed, the accurate vision facilitated by the lens is one of the key adaptations proposed to underlie the evolution

of active predation by ancestral vertebrates, and the subsequent evolutionary success of vertebrates themselves [5, 6]. The unique structural properties of the lens are due to its very high content of long-lived proteins, the crystallins. These derive predominantly from two gene families, the α-crystallin family and the βγ-crystallin family [2]. The structure of βγ-crystallins has been elucidated, and found to have derived from an ancestral protein domain comprised of two symmetrically organised Greek key motifs.

Vertebrates, together with invertebrate urochordates such as the sea squirt *Ciona intestinalis*, comprise phylum Chordata. *C. intestinalis* larvae share a basic chordate body plan with vertebrates, including the possession of a notochord and dorsal neural tube in which an anterior photoreceptor resides in a small brain [7]. Urochordates are, however, thought to have split from the vertebrate lineage prior to the evolution of the lens and the associated co-option of crystallin genes into the visual system. A search of the genome [8] of *C. intestinalis* for βγ-crystallin-like sequences identified one gene coding for a single domain protein with homology to vertebrate βγ-crystallins (for full details of methods, please see supporting online supplementary material). We name this gene *Ci-βγ-crystallin*. Similar searches of the genome of the related urochordate, *Ciona savignyi*, also identified only one homologous gene (http://www.broad.mit.edu). Sequence identity of the *Ciona βγ-crystallins* with vertebrate *βγ-crystallins* was relatively low, and hence to confirm their evolutionary relationship we used X-ray crystallography to solve the structure of Ci-βγ-crystallin protein. The crystal structure of recombinant Ci-βγ-crystallin (Supplementary Tables S1 and S2; Protein Data Bank accession codes for the coordinates and structure factors are 2bv2 and r2bv2sf respectively) shows that the single domain has the

standard βγ-crystallin fold, with two consecutive Greek key motifs, organised about an approximate 2-fold axis (Fig. 1A, B). As with all solved lens βγ-crystallin domains, it has a folded β-hairpin between the first two β-strands of each Greek key motif, a tyrosine corner in the second motif, with the sheet exchanged β-strand of the first motif shorter than in the second. Although two molecules were found in the crystallographic asymmetric unit (Supplementary Table S1), they lack the approximate 2-fold symmetrical pairing typical of vertebrate lens βγ-crystallins. This is consistent with the lack of conservation of interface hydrophobic residues typical of 2-domain lens βγ-crystallins, and the monomeric behaviour of purified Ci-βγ-crystallin. The Ci-βγ-crystallin domain contains two occupied calcium-binding sites that are very similar to those observed in microbial βγ-crystallin domains (Supplementary Table S3). Each calcium-binding site is built from both motifs by virtue of the approximate 2-fold symmetry axis which simultaneously creates two similar binding sites (Fig. 1C). In the crystal lattice of Ci-βγ-crystallin, the protein is bound in layers by the calcium (Supplementary Figure S1). Thus Ci-βγ-crystallin shares structural as well as sequence similarity with vertebrate βγ-crystallins, and in addition has calcium ion binding properties similar to those observed for some non-vertebrate βγ-crystallins and the amphibian protein EDSP [9]. Similar paired calcium-binding βγ-crystallin domain sequences are absent from the known fish and mammalian genomes.

The crystal structure of Ci-βγ-crystallin conclusively confirms its homology to vertebrate βγ-crystallins, and allowed us to construct a structure-based sequence alignment of βγ-crystallin domains (Fig. 2A). In turn, we were able to use this to construct a molecular phylogenetic tree illustrating βγ-crystallin evolution (Fig. 2B).

The *C. intestinalis* sequence is basal to a clade containing vertebrate βγ-crystallins and is related to *G. cydonium* βγ-crystallin. However, the two Greek key motifs of Ci-βγ-crystallin are encoded on separate exons, similar to the organisation of vertebrate *β-crystallins* (Fig. 2B), while the *G. cydonium* gene is an intron-less gene encoding a 2-domain protein [10]. These data show that both *Ci-βγ-crystallin* and the vertebrate *βγ-crystallins* have evolved from a single ancestral gene, encoding a βγ-crystallin domain, which was present in the common ancestor of the chordates.

Analysing the expression of relevant genes from species spanning an evolutionary transition such as the origin of the lens can allow the deduction of the molecular basis for key evolutionary steps. To explore this, we examined the localisation of *Ci-βγ-crystallin* mRNA and protein in *C. intestinalis* embryos, larvae and juveniles by whole mount in-situ hybridisation and immunohistochemistry. *C. intestinalis* has a biphasic life cycle (Fig. 3A). Embryos develop into a swimming larva with a dorsal neural tube and notochord embedded in a muscular tail. In the larval head is a small brain that includes a neuroectodermal sensory vesicle with two sensory organs, the ocellus and the otolith [7], together thought responsible for controlling larval locomotion in the search for a suitable site for metamorphosis [11]. Once located, the larva adheres to the substratum using secretion from three anterior epidermal palps, and subsequently undergoes a radical metamorphosis during which the majority of the brain and tail are reabsorbed. The remaining tissues are extensively remodelled to produce a sedentary adult. The ocellus is a ciliary based photoreceptor system that includes a single pigmented cell, and is considered homologous to the vertebrate retina [3, 4]. In some urochordate larvae, including those of *C. intestinalis*, three cells lie above the pigment cell and, as light must pass through them to reach the

photoreceptors, these are sometimes referred to as lens cells [12, 13]. However, there is no evidence that these cells are homologous to vertebrate lens cells. Similarly, the otolith is not considered homologous to the vertebrate ear [14].

The expression of *Ci-βγ-crystallin* was found to be tightly regulated in a cell specific manner (Fig. 3). *Ci-βγ-crystallin* mRNA was detected in the palps of early and late larvae (Fig. 3B, C; see also www.ghost.kyoto-u.ac.jp). Ci-βγ-crystallin protein was also detected in the palps of larvae (Fig. 3D, E), where the antibody stained the glandular cells, and the protein did not appear to be secreted. Staining of Ci-βγ-crystallin in the palp cells was maintained during attachment and the early part of metamorphosis (Fig. 3G). In later metamorphosis the staining was restricted to a small number of scattered cells (Fig. 3H, I).

In late larvae we also detected Ci-βγ-crystallin in the otolith (Fig. 3E, F). Control pre-immune serum, from the same rabbit as the anti-Ci-βγ-crystallin antibody, did not label this structure. We did not observe staining for mRNA in this cell. This was probably masked by the pigmentation, something we [SMS, unpublished data] and others [15] have observed for other genes. Ci-βγ-crystallin protein in the sensory vesicle was maintained during the early part of metamorphosis, but was not detectable towards the end of metamorphosis. The two pigmented cells of the ascidian sensory vesicle share a common developmental origin, in that they arise from a bilaterally symmetrical pair of cells in the anterior nervous system [16]. These cells have been shown to be initially equivalent, with the potential to form both types of pigment cell. Which forms ocellus and which forms otolith appears to be regulated by Bmp and chordin signalling [17]. Additionally both ocellus and otolith lineages express opsins

[15]. Since anterior photosensory structures are primitive for the bilateria [3, 4], the parsimonious explanation is that both ocellus and otolith evolved from such photosensory structures.

Our data therefore suggest that the chordate βγ-crystallin ancestor was already expressed in a cell-specific manner in derivatives of a primitive visual system prior to the evolution of the lens in the vertebrate lineage. This raises the possibility that the evolution of the lens resulted from the co-option of a pre-existing regulatory circuit also driving the expression of the ancestor of key structural genes, *βγ-crystallins*, in the visual system. An alternative explanation, however, is that βγ-crystallin genes have been independently co-opted in the two lineages. To test these hypotheses, we examined whether the *Ci-βγ-crystallin* promoter region could target expression of a heterologous reporter gene to a vertebrate visual system. First we identified the putative promoter region from the draft genome of *C. intestinalis*. We then cloned this region upstream of GFP to create Ci-βγ-crystallin^{PROM}, electroporated this construct into fertilised *C. intestinalis* eggs, and allowed the resulting transgenic embryos to develop into larvae. Transgenic animals (n=28) showed intense GFP fluorescence in the palps (Fig. 4A, B). A number of these transgenics also showed GFP fluorescence in the ocellus (4/28) or the otolith (5/28), together with occasional fluorescence in cells located adjacent to the ocellus (3/28), which did not appear to include those cells previously referred to as lens cells [12, 13]. Control embryos electroporated with the *Ci-βγ-crystallin* promoter region cloned in reverse orientation in front of GFP (Ci-βγ-crystallin^{REV}) showed no fluorescence. This experiment demonstrates the putative promoter contains the regulatory elements necessary to recapitulate endogenous *Ci-*

βγ-crystallin expression, although it may not contain all the elements necessary to repress ocellus expression.

Next, we introduced Ci-βγ-crystallin^{PROM} into *Xenopus laevis*. In the resulting transgenic *X. laevis* tadpoles (n=24), the *Ci-βγ-crystallin* promoter specifically directed GFP expression to the developing vertebrate visual system, including the optic tectum, the optic nerve/retinal ganglion cells and the lens (67%, 46% and 54% of animals respectively) (Fig. 4C-F). Faint expression was also occasionally seen in the otic vesicle and nasal epithelium (20% and 25% of animals respectively). *X. laevis* embryos transgenic for control constructs, including Ci-βγ-crystallin^{REV} and one derived from the *C. intestinalis Brachyury* promoter [18] linked to GFP, did not show similar visual-system specific expression (data not shown). This experiment confirms that the regulatory circuitry driving *βγ-crystallin* gene expression in the visual system is conserved between *C. intestinalis* and *X. laevis*. Vertebrate eyes have a binary origin, with the retina, optic nerve and optic tectum arising from the central nervous system, while the lens arises from an ectodermal placode. Notably, the *Ci-βγ-crystallin* promoter targets expression to both lens (where vertebrate βγ-crystallins are primarily expressed) and neural components. This indicates the *Ci-βγ-crystallin* promoter is probably recognising a visual-system wide regulatory cue, not a lens-specific cue. A degree of higher-level regulatory similarity between eyes in different taxa, including the vertebrate lens and neural visual system, has been previously recognised [19]. We speculate that such conserved transcription factors may be responsible for the observed pattern of *Ci-βγ-crystallin* promoter activation.

Our study demonstrates that the vertebrate *βγ-crystallin* genes have evolved from a single ancestral gene present in the common ancestor of the chordates, and that this gene also gave rise to the single *βγ-crystallin* ortholog in *C. intestinalis*. While our data do not exclude the possibility that the lens itself evolved earlier than currently thought, and has degenerated in modern urochordates, there is no evidence to support this view. Hence, we propose that this ancestral gene was already expressed in the neurectodermal visual system prior to the evolution of the lens, and that its regulation was conserved during the evolution of the ectodermal lens. We therefore conclude that the evolution of the lens did not derive from a new association between a visual system regulatory circuit and co-opted lens structural genes, but from the re-use of a pre-existing regulatory interaction linking these components in the central nervous system of a primitive chordate.

Acknowledgments

NHL thanks Anita Kaan for her help in raising antibodies against recombinant Ci-βγ-crystallin, Tim Kortum for help searching the database, Gerard Martens for providing the *Xenopus* facilities and the National Institute of Genetics (Japan) for providing a *Ci-βγ-crystallin* cDNA clone. We thank Ole Madsen for his help with molecular phylogenetic analysis. X-ray diffraction data were collected at the ESRF, Grenoble. AGP, OAB and CS are grateful for the financial support of the Medical Research Council, UK. SMS thanks the BBSRC for financial support, and the staff of Sparkes Yacht Haven and Northney Marina for allowing the collection of ascidians from their pontoons.

References

1. Land, M.F., and Nilsson, D.-E. (2002). Animal Eyes (Oxford: Oxford University press).

2. Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N.H., Slingsby, C., and Tardieu, A. (2004). Ageing and vision: structure, stability and function of lens crystallins. Prog Biophys Mol Biol *86*, 407-485.

3. Arendt, D., and Wittbrodt, J. (2001).Reconstructing the eyes of Urbilateria. Phil Trans R Soc Lond B *356*, 1545-1563.

4. Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A.W., and Wittbrodt, J. (2004). Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. Science *306*, 869-871.

5. Shimeld, S.M., and Holland, P.W.H. (2000). Vertebrate innovations. Proc Natl Acad Sci USA *97*, 4449-4452.

6. Gans, C., and Northcutt, R.G. (1983). Neural crest and the origin of vertebrates: a new head. Science *220*, 268-273.

7. Satoh, N. (1994). Developmental Biology of Ascidians (Cambridge: Campridge University Press).

8. Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., Harafuji, N., Hastings, K.E., Ho, I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, I.A., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H.G., Awazu, S., Azumi, K., Boore, J., Branno, M., Chin-Bow, S., DeSantis, R., Doyle, S., Francino, P., Keys, D.N., Haga, S., Hayashi, H., Hino, K., Imai, K.S., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B.I., Makabe, K.W., Manohar, C., Matassi, G., Medina, M., Mochizuki, Y., Mount, S., Morishita,

T., Miura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oishi, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Shin-i, T., Spagnuolo, A., Stainier, D., Suzuki, M.M., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wada, S., Zhang, C., Hyatt, P.D., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., and Rokhsar, D.S. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science *298*, 2157-2167.

9. Ogawa, M., Takabatake, T., Takahashi, T.C., and Takeshima, K. (1997). Metamorphic change in EP37 expression: members of the betagamma-crystallin superfamily in newt. Dev Genes Evol *206*, 417-424.

10. Di Maro, A., Pizzo, E., Cubellis, M.V., and D'Alessio, G. (2002). An intron-less betagamma-crystallin-type gene from the sponge *Geodia cydonium*. Gene *299*, 79-82.

11. Sakurai, D., Goda, M., Kohmura, Y., Horie, T., Iwamoto, H., Ohtsuki, H., and Tsuda, M. (2004). The role of pigment cells in the brain of ascidian larva. J Comp Neurol *475*, 70-82.

12. Dilly, P.N. (1964). Studies on receptors in the cerebral vesicle of the ascidian tadpole. 2. The ocellus. Q. J. Micros. Sci. *105*, 13-20.

13. Ohtsuki, H. (1990). Statocyte and ocellar pigment cell in embryos and larvae of the ascidian, *Styela plicata* (Leseur). Develop Growth Differ *1*, 85-90.

14. Mazet, F., Hutt, J.A., Millard, J., Graham, A., and Shimeld, S.M. (2005). Molecular evidence from *Ciona intestinalis* for the evolutionary origin of vertebrate cranial placodes. Dev Biol *282*, 494-508.

15. Nakashima, Y., Kusakabe, T., Kusakabe, R., Terakita, A., Shichida, Y., and Tsuda, M. (2003). Origin of the vertebrate visual cycle: genes encoding retinal photoisomerase and two putative visual cycle proteins are expresssed in whole brain of a primitive chordate. J Comp Neurol *460*, 180-190.

16. Taniguchi, K., and Nishida, H. (2004). Tracing cell fate in brain formation during embryogenesis of the ascidian *Halocynthia roretzi*. Dev Growth Differ *46*, 163-180.

17. Darras, S., and Nishida, H. (2001). The BMP/CHORDIN antagonism controls sensory pigment cell specification and differentiation in the ascidian embryo. Dev Biol *236*, 271-288.

18. Corbo, J.C., Levine, M., and Zeller, R.W. (1997). Characterization of a notochotd-specific enhancer from the *Brachyury* promoter region of the ascidian, *Ciona intestinalis*. Development *124*, 589-602.

19. Gehring, W.J. (2005). New perspectives on eye development and the evolution of eyes and photoreceptors. J Hered *96*, 171-184.

20. Krasko, A., Muller, I.M., and Muller, W.E. (1997). Evolutionary relationships of the metazoan beta gamma-crystallins, including that from the marine sponge *Geodia cydonium*. Proc R Soc Lond B Biol Sci *264*, 1077-1084.

21. Wistow, G., Summers, L., and Blundell, T. (1985). *Myxococcus xanthus* spore coat protein S may have a similar structure to vertebrate lens beta gamma-crystallins. Nature *315*, 771-773.

22. Wistow, G. (1990). Evolution of a protein superfamily: relationships between vertebrate lens crystallins and microorganism dormancy proteins. J Mol Evol *30*, 140-145.

23. Clout, N.J., Kretschmar, M., Jaenicke, R., and 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 (Camb) *9*, 115-124.

24. Chan, C.W., Saimi, Y., and Kung, C. (1999). A new multigene family encoding calcium-dependent calmodulin-binding membrane proteins of *Paramecium tetraurelia*. Gene *231*, 21-32.

25. Ray, M.E., Wistow, G., Su, Y.A., Meltzer, P.S., and Trent, J.M. (1997). AIM1, a novel non-lens member of the betagamma-crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. Proc Natl Acad Sci U S A *94*, 3229-3234.

Figure Legends

Figure 1

The sequence and crystal structure of Ci-βγ-crystallin.

(A) The ribbon diagram shows the first Greek key motif (turquoise) pairing with the second motif (yellow) and the two calcium ions are shown as red spheres. The rms deviation between Ci-βγ-crystallin and a human gamma-crystallin D domain for main chain atoms is 1.1 Angstrom.

(B) The protein sequence, translated from three exons of the DNA sequence from 17637 to 18636 in Contig 0605, is shown with the two motifs aligned and colour coded to match the structure. The first exon sequence (the methionine is cleaved off in the recombinant protein) is shown in bold. The region encoded by first exon and that encoded by the third exon is underlined. The phase of the exon/intron junction is 0 for the first intron and 2 for the second intron. This genomic organisation, including the phases of the exon/intron junctions is identical to that of the (first half) of the mammalian β-crystallin genes.

(C) Each calcium-binding site is formed from two main chain oxygen atoms and one side chain atom (SER OG) from the same motif and one side chain atom (ASP OD1) from the partner motif. In the alignment these calcium-binding residues are highlighted in dark blue and red, while the two residues in each motif that bind calcium with their side chains are also in bold. The similar backbone and conserved side chains of each Ci-βγ-crystallin calcium-binding site are shown in the detailed structure, with the backbone in brown and the calcium binding regions in green. The residues involved from the first motif are labelled in blue and from the second in orange. Each calcium-binding site thus requires a subset of residues from both motifs, whereas domain paired calcium-binding sites need all the sequence-highlighted

residues. The glutamate side chain shown in blue comes from a symmetry-related molecule in the lattice and contributes to the formation of the calcium bound lattice layers (Supplementary Figure S1).

Figure 2

(A) Structure-based sequence alignment of Ci-βγ-crystallin and *C. savignyi* βγ-crystallin with βγ-crystallin domains from a selection of vertebrate βγ-crystallins and from microbial orthologs. The A type motif is shown above and B type motif below. The secondary structure of Ci-βγ-crystallin is indicated at the bottom of each motif alignment: beta sheet strands as arrows, helices as cylinders. Calcium binding residues are colour coded as follows: residues providing main chain atoms are in pink, side chain atoms are in green and both types are in red. Numbers at the start and end of each sequence refer to the start and end position respectively of the motif in the respective protein. Species, gene abbreviations and associated references are: *Geodia: Geodia cydonium* βγ-crystallin [20]. *Cynops: Cynops pyrrhogaster* GEP [9]. Human: *Homo sapiens*. Protein S: *Myxococcus xanthus* spore coat protein S [21]. Spherulin: spherulin 3A from *Physarum polycephalum* [22]. Cholera: *Vibrio cholerae* [23]. *Paramecium: Paramecium tetraurelia* [24]. Q8N7F1 is the TREMBL identifier for an additional human βγ-crystallin homologue.

(B) Molecular phylogeny of βγ-crystallin domains. Maximum likelihood tree based on the structure-based sequence alignment of Ci-βγ-crystallin with other βγ-crystallins. The model used is JTT+I+G8. Branch lengths are not shown, as the sequences are too distant for branch lengths to be reliable. Figures adjacent to nodes indicate percentage bootstrap support, and only values greater than 70% are shown.

For the human β- and γ-crystallins, the human AIM1 protein as well as for Ci-βγ-crystallin and the *G. cydonium* βγ-crystallin the gene structure of the domain coding regions is shown schematically, where the boxed M indicates a motif coding region. Of the 12 βγ-crystallin like motif encoding exons in the human AIM1 gene, only two are shown. Note that each motif of Ci-βγ-crystallin is encoded by a separate exon as in the vertebrate β-crystallin genes and the AIM1 genes [25]. In the vertebrate γ-crystallin genes, an exon encodes two motifs, while the *G. cydonium* βγ-crystallin gene is intron-less.

Figure 3

Localisation of *Ci-βγ-crystallin* expression in *C. intestinalis* larvae and metamorphs. **(A)** schematic life cycle of *C. intestinalis*. A planktonic embryo produces a motile larva which seeks a suitable site to settle and metamorphose. Initial attachment is via the anterior palps, while metamorphosis involves extensive remodelling of the body, including loss of the tail and most of the central nervous system. **(B)** recently hatched larvae, showing expression of *Ci-βγ-crystallin* mRNA in the palps (p). **(C)** slightly older larva than shown in (B). ot, otolith. oc, ocellus. **(D)** early larva showing localisation of Ci-βγ-crystallin protein (red) in the palps. **(E)** late larva showing localisation of Ci-βγ-crystallin protein in the palps and otolith cell. **(F)** larva immediately before settlement, showing localisation of Ci-βγ-crystallin protein in the palps and otolith cell. **(G)** Metamorph shortly after the initiation of metamorphosis. The tail has contracted (asterisk). *Ci-βγ-crystallin* protein is located in the former palp cells (arrow) and the otolith cell. **(H)** metamorphs in which the stalk that connects the animal to the substrate has begun to elongate. *Ci-βγ-crystallin* protein is still weakly

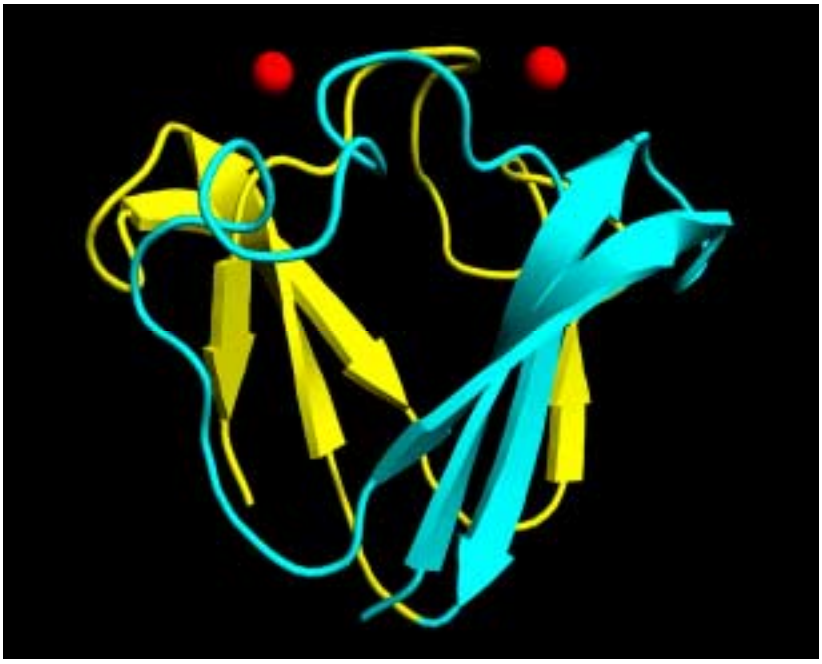
detected in the otolith cell, and in the remnants of the palps (arrow). **(I)** In older metamorphs, the Ci-βγ-crystallin positive palp cells (arrow) begin to disperse.

Figure 4

The *Ci-βγ-crystallin* upstream region driving expression of GFP in transgenic *Ciona intestinalis* larvae and *Xenopus laevis* tadpoles. (**A**) Larva electroporated with Ci-βγ-crystallin^{PROM} viewed from the right lateral aspect. GFP fluorescence is visible in one palp (p, out of focus), the otolith (ot) and in a cell adjacent to the ocellus (oc). (**B**) A different larva transgenic for the same construct seen in left lateral aspect, showing expression in one palp and in the ocellus. The pigmented cells of the otolith and ocellus derive from the same lineage on either side of the midline [16, 17], and consequently ocellus expression may reflect this shared developmental history (**C-F**) *X. laevis* tadpoles transgenic for Ci-βγ-crystallin^{PROM}. In the majority of tadpoles expression of GFP was restricted to varying combinations of the following tissues: brain (most prominent in midbrain) (**C, D**), lens (**C-E**) optic nerve (**D, E**), and otic vesicle (arrowheads in **F**). l, lens; nf, non-specific fluorescence in liver and yolk; on, optic nerve; ot, optic tectum.

Figure 1

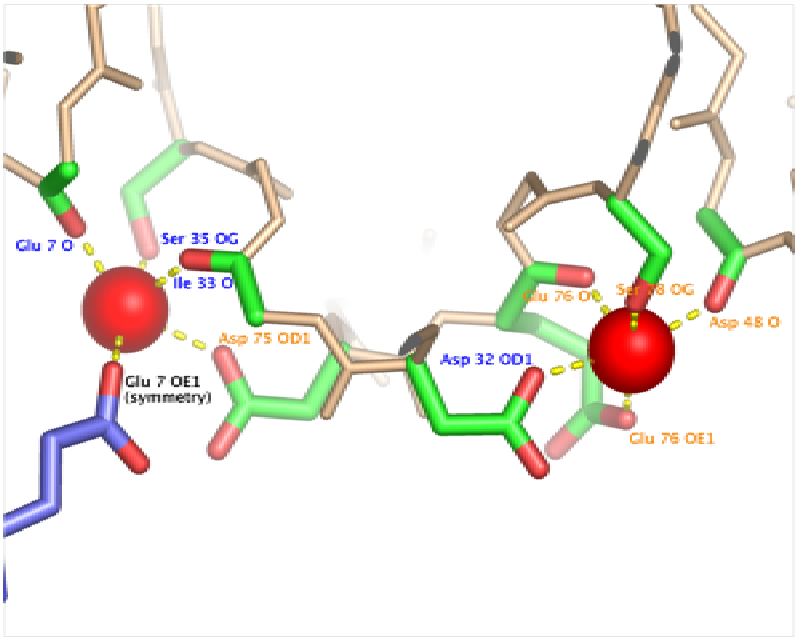
A



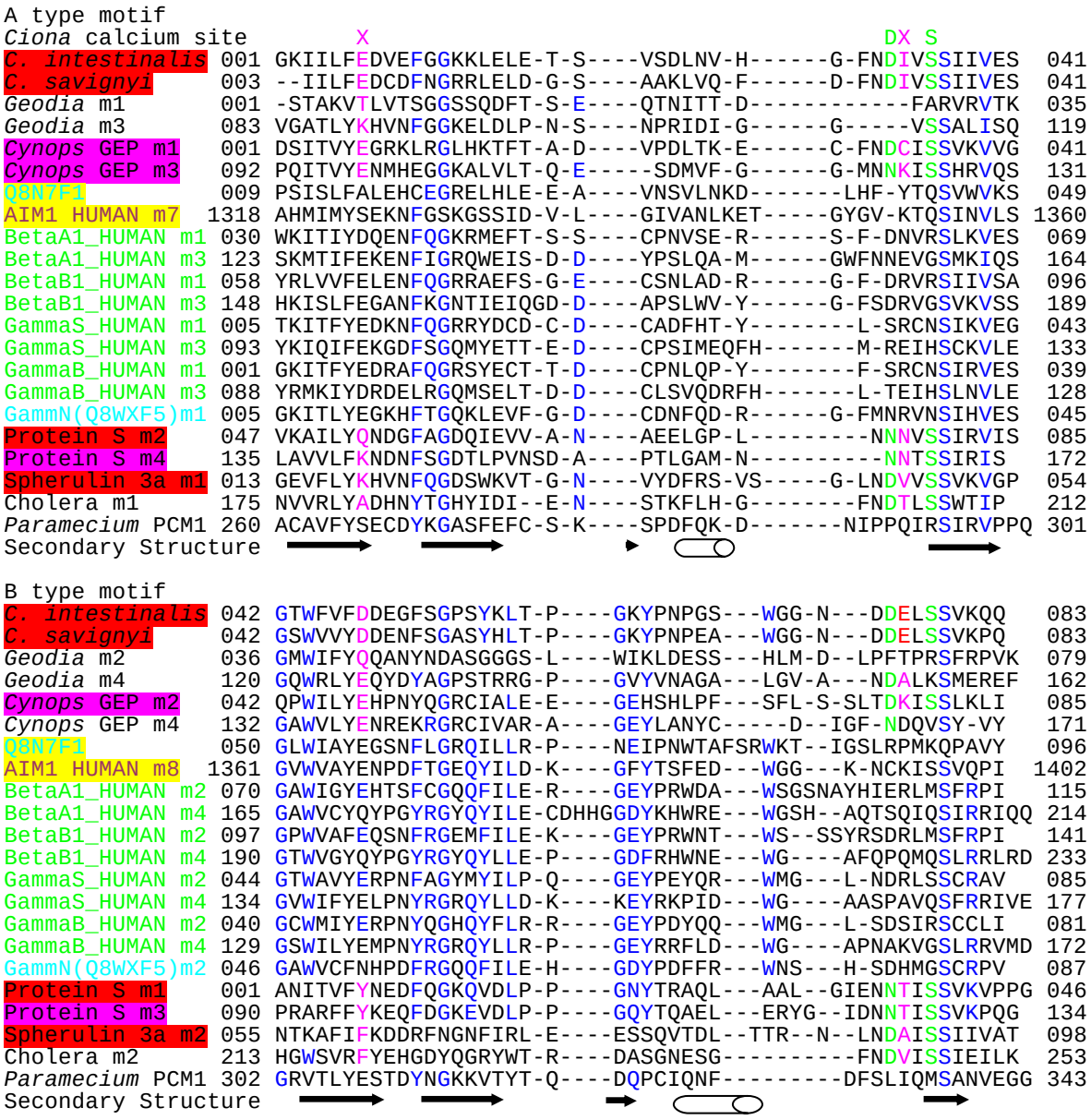
B

```
1    MGKIILFEDVEFGGKKLELET-SVSDLNVH-GFNDIVSSIIVES 41
42     GTWFVFDDEGFSGPSYKLTPGKYPNPGSWGGNDDELSSVK-QQ 83
```

C



Figure


A type motif
Ciona calcium site X DX S
C. intestinalis 001 GKIILFEDVEFGGKKLELE-T-S----VSDLNV-H------G-FNDIVSSIIVES 041
C. savignyi 003 --IILFEDCDFNGRRLELD-G-S----AAKLVQ-F------D-FNDIVSSIIVES 041
Geodia m1 001 -STAKVTLVTSGGSSQDFT-S-E----QTNITT-D-----------FARVRVTK 035
Geodia m3 083 VGATLYKHVNFGGKELDLP-N-S----NPRIDI-G------G-----VSSALISQ 119
Cynops GEP m1 001 DSITVYEGRKLRGLHKTFT-A-D----VPDLTK-E------C-FNDCISSVKVVG 041
Cynops GEP m3 092 PQITVYENMHEGGKALVLT-Q-E-----SDMVF-G------G-MNNKISSHRVQS 131
Q8N7F1 009 PSISLFALEHCEGRELHLE-E-A----VNSVLNKD-------LHF-YTQSVWVKS 049
AIM1_HUMAN m7 1318 AHMIMYSEKNFGSKGSSID-V-L----GIVANLKET-----GYGV-KTQSINVLS 1360
BetaA1_HUMAN m1 030 WKITIYDQENFQGKRMEFT-S-S----CPNVSE-R------S-F-DNVRSLKVES 069
BetaA1_HUMAN m3 123 SKMTIFEKENFIGRQWEIS-D-D----YPSLQA-M------GWFNNEVGSMKIQS 164
BetaB1_HUMAN m1 058 YRLVVFELENFQGRRAEFS-G-E----CSNLAD-R------G-F-DRVRSIIVSA 096
BetaB1_HUMAN m3 148 HKISLFEGANFKGNTIEIQGD-D----APSLWV-Y------G-FSDRVGSVKVSS 189
GammaS_HUMAN m1 005 TKITFYEDKNFQGRRYDCD-C-D----CADFHT-Y--------L-SRCNSIKVEG 043
GammaS_HUMAN m3 093 YKIQIFEKGDFSGQMYETT-E-D----CPSIMEQFH-------M-REIHSCKVLE 133
GammaB_HUMAN m1 001 GKITFYEDRAFQGRSYECT-T-D----CPNLQP-Y--------F-SRCNSIRVES 039
GammaB_HUMAN m3 088 YRMKIYDRDELRGQMSELT-D-D----CLSVQDRFH-------L-TEIHSLNVLE 128
GammN(Q8WXF5)m1 005 GKITLYEGKHFTGQKLEVF-G-D----CDNFQD-R------G-FMNRVNSIHVES 045
Protein S m2 047 VKAILYQNDGFAGDQIEVV-A-N----AEELGP-L---------NNNVSSIRVIS 085
Protein S m4 135 LAVVLFKNDNFSGDTLPVNSD-A----PTLGAM-N---------NNTSSIRIS 172
Spherulin 3a m1 013 GEVFLYKHVNFQGDSWKVT-G-N----VYDFRS-VS-----G-LNDVVSSVKVGP 054
Cholera m1 175 NVVRLYADHNYTGHYIDI--E-N----STKFLH-G--------FNDTLSSWTIP 212
Paramecium PCM1 260 ACAVFYSECDYKGASFEFC-S-K----SPDFQK-D-------NIPPQIRSIRVPPQ 301
Secondary Structure
B type motif
C. intestinalis 042 GTWFVFDDEGFSGPSYKLT-P----GKYPNPGS---WGG-N---DDELSSVKQQ 083
C. savignyi 042 GSWVVYDDENFSGASYHLT-P----GKYPNPEA---WGG-N---DDELSSVKPQ 083
Geodia m2 036 GMWIFYQQANYNDASGGGS-L----WIKLDESS---HLM-D--LPFTPRSFRPVK 079
Geodia m4 120 GQWRLYEQYDYAGPSTRRG-P----GVYVNAGA---LGV-A---NDALKSMEREF 162
Cynops GEP m2 042 QPWILYEHPNYQGRCIALE-E----GEHSHLPF---SFL-S-SLTDKISSLKLI 085
Cynops GEP m4 132 GAWVLYENREKRGRCIVAR-A----GEYLANYC-----D--IGF-NDQVSY-VY 171
Q8N7F1 050 GLWIAYEGSNFLGRQILLR-P----NEIPNWTAFSRWKT--IGSLRPMKQPAVY 096
AIM1_HUMAN m8 1361 GVWVAYENPDFTGEQYILD-K----GFYTSFED---WGG---K-NCKISSVQPI 1402
BetaA1_HUMAN m2 070 GAWIGYEHTSFCGQQFILE-R----GEYPRWDA---WSGSNAYHIERLMSFRPI 115
BetaA1_HUMAN m4 165 GAWVCYQYPGYRGYQYILE-CDHHGGDYKHWRE---WGSH--AQTSQIQSIRRIQQ 214
BetaB1_HUMAN m2 097 GPWVAFEQSNFRGEMFILE-K----GEYPRWNT---WS--SSYRSDRLMSFRPI 141
BetaB1_HUMAN m4 190 GTWVGYQYPGYRGYQYLLE-P----GDFRHWNE---WG----AFQPQMQSLRRLRD 233
GammaS_HUMAN m2 044 GTWAVYERPNFAGYMYILP-Q----GEYPEYQR---WMG---L-NDRLSSCRAV 085
GammaS_HUMAN m4 134 GVWIFYELPNYRGRQYLLD-K----KEYRKPID---WG----AASPAVQSFRRIVE 177
GammaB_HUMAN m2 040 GCWMIYERPNYQGHQYFLR-R----GEYPDYQQ---WMG---L-SDSIRSCCLI 081
GammaB_HUMAN m4 129 GSWILYEMPNYRGRQYLLR-P----GEYRRFLD---WG----APNAKVGSLRRVMD 172
GammN(Q8WXF5)m2 046 GAWVCFNHPDFRGQQFILE-H----GDYPDFFR---WNS---H-SDHMGSCRPV 087
Protein S m1 001 ANITVFYNEDFQGKQVDLP-P----GNYTRAQL---AAL--GIENNTISSVKVPPG 046
Protein S m3 090 PRARFFYKEQFDGKEVDLP-P----GQYTQAEL---ERYG--IDNNTISSVKPQG 134
Spherulin 3a m2 055 NTKAFIFKDDRFNGNFIRL-E----ESSQVTDL--TTR--N--LNDAISSIIVAT 098
Cholera m2 213 HGWSVRFYEHGDYQGRYWT-R----DASGNESG----------FNDVISSIEILK 253
Paramecium PCM1 302 GRVTLYESTDYNGKKVTYT-Q----DQPCIQNF---------DFSLIQMSANVEGG 343
Secondary Structure

Figure

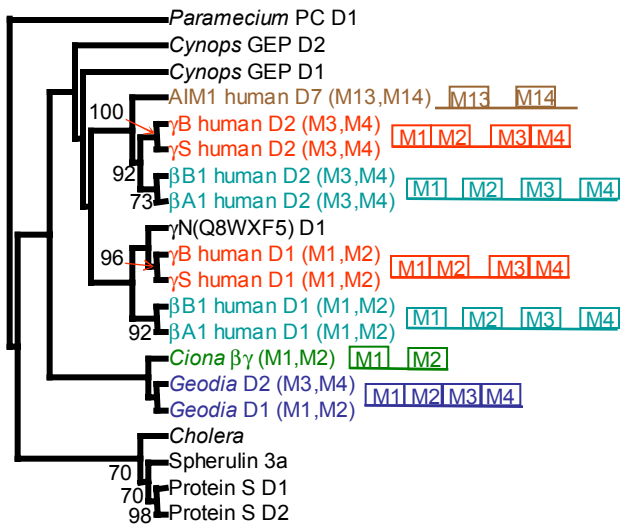


Figure 2B

Figure

Click here to download high resolution image

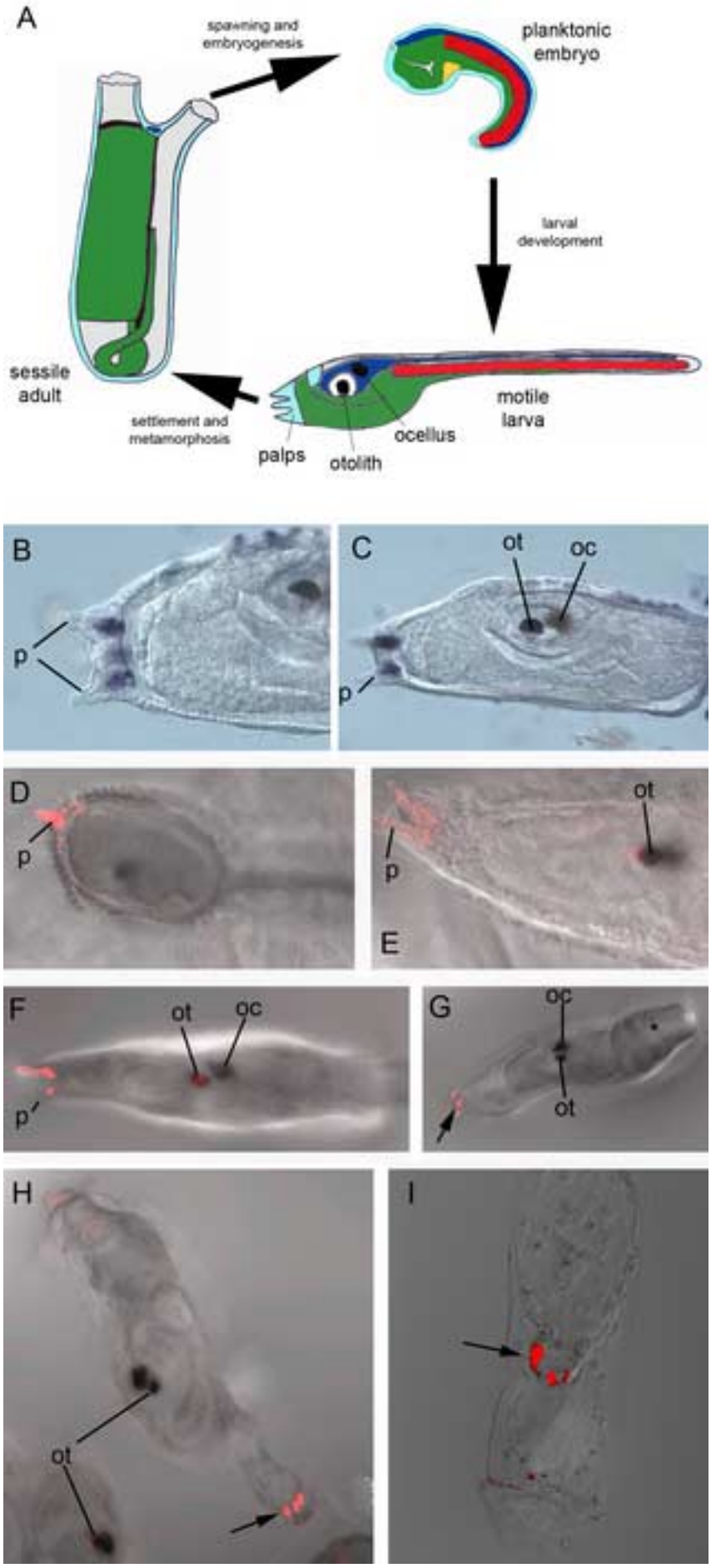
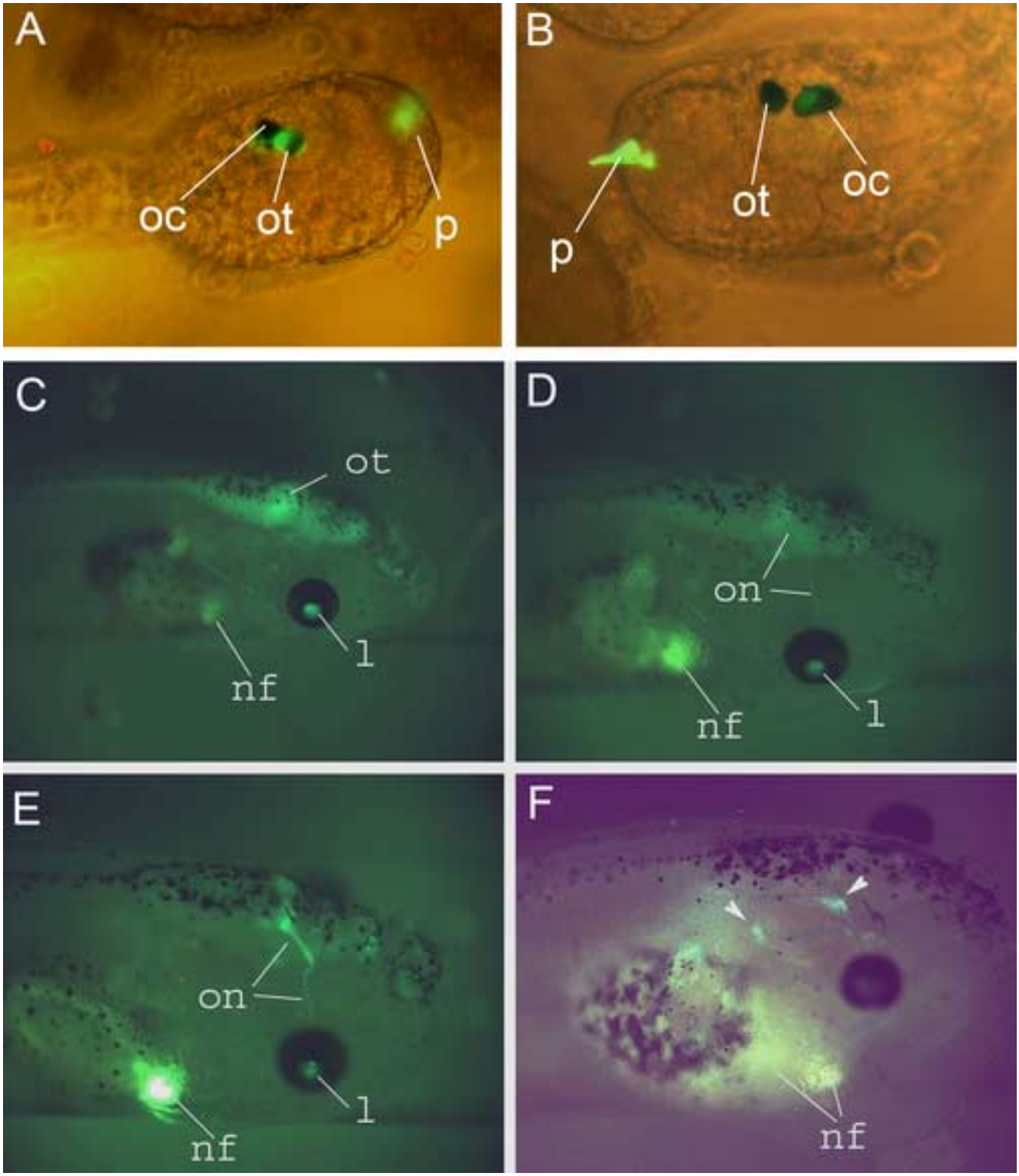


Figure
Click here to download high resolution image



Urochordate βγ-crystallin and the evolutionary origin of the vertebrate eye lens.

Sebastian M. Shimeld,[1] Andrew G. Purkiss,[2] Ron P.H. Dirks,[3] Orval A. Bateman,[2] Christine Slingsby[2] and Nicolette H. Lubsen[3].

1. Department of Zoology, University of Oxford, The Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK

2. Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

3. Department of Biochemistry, Faculty of Science, Radboud University Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

Corresponding author: Sebastian Shimeld
Email: Sebastian.Shimeld@zoo.ox.ac.uk

Supporting Online Supplementary Material

Contents

Experimental procedures

Cloning and Recombinant DNA Methods

The ~1-kb intergenic region of the head-to-head-oriented *cubilin* and *Ci-βγ-crystallin* genes (scaffold 605; www.jgi.doe.gov) was amplified by the polymerase chain reaction (PCR) from 10 ng *Ciona intestinalis* DNA using sense primer 5'-AGC TGT CGA CTA ATT CTT ACT GTT CGG TTG AAA CTC-3' and antisense primer 5'-AGC TAA GCT TGA AAC TTC GAT TGT ACA AAA TGC G-3' (35 cycles: 30'', 95°C; 30'', 55°C; 2', 72°C; Fast start high fidelity kit, Roche, Germany). The resulting PCR fragment was cloned into the *Sal*I and *Hin*dIII sites of the *Xenopus* expression vector pCS2+ [1], yielding pCiCrys2+. The coding sequence for enhanced green fluorescent protein (EGFP) was excised from pIRES2-EGFP (BD Biosciences Clontech, USA) using *Msc*I and *Not*I and, via an intermediary step in pBluescript-SK(-) (Stratagene, USA), ligated into the *Eco*RI and *Stu*I sites of pCiCrys2+, resulting in pCiCrys-GFP (Ci-βγ-crystallin^{PROM}). The control plasmid pCiCrys(rev)-GFP (Ci-βγ-crystallin^{REV}) was generated by reamplifying the intergenic region from pCiCrys-

GFP using sense primer 5'-AGC TGT CGA CGA AAC TTC GAT TGT ACA AAA TGC G-3' and antisense primer 5'-AGC TAA GCT TTA ATT CTT ACT GTT CGG TTG AAA CTC-3' (35 cycles: 30'', 95°C; 30'', 55°C; 2', 72°C) and cloning it, in the reverse orientation, into the *Sal*I and *Hin*dIII sites of pCiCrys-GFP.

Stable transgenesis of *Xenopus laevis*

A *Sal*I-*Not*I fragment, containing the Ci-βγ-crystallin intergenic region in the forward or reverse orientation, the EGFP open reading frame and the SV40 polyadenylation signal, was gel purified from the pCS2+ vector backbone from Ci-βγ-crystallin^{PROM} and Ci-βγ-crystallin^{REV} respectively and recovered using the GFX gel band purification kit (Amersham, UK). Transgenesis of *Xenopus laevis* was performed according to Kroll and Amaya [2], with modifications [3]. In summary: 250,000 sperm nuclei were mixed with ~200 ng DNA fragment, incubated for 15 min at room temperature and diluted in 500 µl sperm dilution buffer (250 mM sucrose, 75 mM KCl, 0.5 mM spermidine trihydrochloride, 0.2 mM spermidine tetrahydrochloride, 5 mM $MgCl_2$, pH 7.4). Eggs were dejellied in 2% cystein/1 x MMR (1 x MMR: 0.1 M NaCl, 0.02 M KCl, 0.01 M $MgCl_2$, 0.015 M $CaCl_2$ en 0.5 M HEPES pH 7.5), transferred to 6% Ficoll/0.4 x MMR and injected with 10 nl of the diluted nuclei/DNA mixture at 17°C. At the 4-cell stage, the embryos were transferred to 6% Ficoll/0.1 x MMR and incubated overnight at 17°C. At the gastrula stage, the embryos were transferred to 0.1 x MMR and incubated at 22°C. EGFP-positive tadpoles were photographed at stage 45 [4], using a MZ FLIII fluorescence stereomicroscope provided with a DC200 camera (Leica microsystems, Switzerland).

Expression and purification of recombinant Ci-βγ-crystallin

The *Ci-βγ-crystallin* cDNA clone (cilv010f04) was obtained from the National Genetics Institute (Japan). The cDNA sequence was excised by PCR, simultaneously creating a NdeI site comprising the start codon and a BamHI site 3'of the stop codon. The PCR product was cloned in pGEM-T Easy (Promega) and sequence verified. The cDNA sequence was then cloned NdeI/BamHI in pET3a. The pET3a recombinant was transformed into BL21(DE3) strain (Novagen), grown and induced as described previously for lens β-crystallins [5]. Following harvesting by centrifugation, cells were resuspended in Bugbuster (Novagen) and Pefabloc SC (Merck). Cells were lysed by two passes through an Emulsiflex Homogeniser (Glen Creston Ltd), followed by addition of DNAse I and sonication (4 x 15 seconds) with cooling. The lysate was centrifuged at 18000 rpm at 4°C for 20 minutes, then the supernatant was dialysed overnight at 4°C using 6 kDa MWCO dialysis tubing against 25 mM Tris-HCl, pH 7.5, 2 mM DTT (Buffer A). The dialysed lysate was then passed through 0.45 and then 0.2 mm filters before loading onto an ion exchange column (HiPrep 16/10 Q FF, Amersham Biosciences) attached to an AKTA purifier (Amersham Biosciences). Analysis of the peaks by SDS PAGE showed that the protein eluted in a broad peak around 30% Buffer B (Buffer A/1 M NaCl). A sample of the peaks was desalted using a BioSpin column (BioRad), the mass measured using the Micromass ESMS and found to agree with the calculated sequence mass without the N-terminal methionine. The peaks from several runs were collected, concentrated in an Amicon Ultrafiltration cell fitted with a YM3 membrane and then loaded onto a HiPrep 26/10 desalting column run in Buffer A. The eluted protein peak was then loaded onto a Mono Q 10/100 column run in Buffer A with a linear gradient of Buffer B, the Ci-βγ-crystallin peak eluting at 20 % Buffer B. The final polishing step involved concentrating the Mono Q peaks using the Amicon Ultrafiltration cell with YM3 membrane and then loading the protein solution onto a Sephacryl S300HR run in 50 mM MES, 200 mM NaCl, pH 6.05 (Buffer C).

Crystallisation of Ci-βγ-crystallin

Protein concentration was estimated from absorption at 280 nm based on an extinction coefficient of 1.5 for a 1 mg/ml solution. Protein was initially concentrated in Buffer C to around 20 mg/ml. A large number of trials were made which included desalting, treating with EDTA and adding calcium acetate. Crystals grew without adding calcium but with better morphology when calcium was added. Crystals were grown by vapour diffusion at 4°C from 1 µl of protein solution and 1 µl of reservoir solution, and equilibrated against 1 ml of reservoir solution. Diffraction datasets were collected from 2 crystals grown under slightly different conditions.

Crystal 1: The protein was in buffer C at 8 mg/ml. The reservoir solution was 0.2 M $(NH_4)_2SO_4$, 1 M sodium acetate pH 4.6, 30% PEGmme 2000, with 10 µl of hexane-1,6 diol added.

Crystal 2: The protein was desalted and concentrated to around 8 mg/ml in 100 mM sodium acetate, pH 5.5. The reservoir solution was 0.2 M $(NH_4)_2SO_4$, 0.35 M sodium acetate pH 4.6, 20% PEGmme, 1 mM calcium acetate.

As crystals were stacks of square or hexagonal plates which proved difficult to separate, individual stacks were frozen, using 30% glycerol made up in the well solution, as cryoprotectant. Crystals were soaked for one to two minutes before being flash frozen in liquid nitrogen.

Estimation of solution molecular weight

The molecular weight of the protein at 8 mg/ml in Buffer C was evaluated by dynamic light scattering on a dp801 dls instrument (Protein Solutions). The average over 15 readings gave a diffusion coefficient D_T of 1350 x 10^{-13} m2/s showing that the protein was monomeric in solution when calibrated against protein standards. The data was of the highest standard with virtually all baseline values within the range 1.000+/- 0.001. Almost all SOS values were below 5 and the majority were below 2 showing that the quality of the data was statistically valid.

X-ray diffraction analysis

A 90° diffraction dataset was collected from crystal 1 on the in-house Rigaku RU-H3R rotating anode X-ray source with 450 s per 1^O oscillation. A crystal to detector distance of 150 mm gave diffraction to 2.0 Å. A 180° diffraction dataset on crystal 2 was collected at ESRF beamline ID14-2, with 1^O oscillation and 12 s exposure per frame. A crystal to detector distance of 125 mm gave diffraction to 1.50 Å. A second dataset was collected from this crystal at a low-resolution (180^O with 1^O oscillation with 4 s exposure per frame, detector at 305 mm).

Structure solution and refinement

The data from crystal 1 were processed using MOSFLM [6]. The superimposed lattices from the stacked crystals were dealt with by deselecting spots found by the spot search before autoindexing. The data were scaled with the CCP4 program [7] SCALA followed by TRUNCATE. Phases were obtained using molecular replacement with the program PHASER [8] (version 1.1). As vertebrate lens crystallins have two domains per polypeptide chain, an ensemble of the C-terminal domains of three crystallins was used as the search model. The C-terminal domains from human βB1-crystallin (pdb code: 1OKI) (40% identity used as input for PHASER), human γD-crystallin (1HK0) (30% identity) and human γS-crystallin (1HA4) (30% identity), were superimposed using backbone atoms with MOLMOL [9]. The molecular replacement showed two molecules in the asymmetric unit, with the first having a log likelihood score of 55.83 and sigma of 7.05 and the second (with the first molecule fixed) having a log likelihood score of 181.66 and a sigma of 12.41. The human γS-crystallin C-terminal domain was used as the initial model for model

building, using ARP_WARP[10], following rigid body refinement with REFMAC[11]. Following manual rebuilding, addition of water molecules and refinement using XFIT [12] and ARP_WARP and REFMAC it became clear that several assigned water sites were metal ions. Examination of these sites showed that they were likely to be calcium ions, both from the size of the positive peak in weighted difference maps and the similarity of the orientation of the ligands when compared to spherulin 3a (pdb code: 1HDF).
The dataset from crystal 2 was also processed using MOSFLM, although autoindexing proved difficult and approximate cell dimensions from the first crystal had to be used as a starting point for processing. The data were scaled as before and molecular replacement undertaken using MOLREP[13], with the refined solution from the first crystal used as search model. The first solution had R = 0.527, correlation = 0.384, the second solution (with the first fixed) had R = 0.467, correlation = 0.511. Refinement proceeded, with rigid-body refinement using REFMAC, automated rebuilding of the structure using ARP_WARP, docking of the sequence using GUISIDE, then manual rebuilding, refinement and addition of calciums, waters and counterions with XFIT, REFMAC and ARP_WARP. During refinement, the resolution was cut-off to 1.55Å, the point where I/sigI fell below 2.0.

Production of transgenic *C. intestinalis* embryos

Transgenic *C. intestinalis* embryos were made essentially as described by Corbo et al [14]. Briefly, oocytes were dissected from adult gonoducts, fertilised in vitro and immediatley chemically dechorionated [15]. 200 µl of eggs (typically between 200 and 500 individual eggs) were gently mixed with 500 µl 0.77 M mannitol containing 50 µg of circular plasmid DNA, then transferred to a 0.4 cm electroporation cuvette. The cuvette was then pulsed once (50 V, 16 milliseconds) in a BTX Electrosquare porator T820, after which the eggs were transferred to an agarose-sea water petri dish and allowed to develop overnight at 18ºC, by which time they had reached the larval stage. Surviving larvae were viewed on a Zeiss Axioskop II equipped with fluorescence and a GFP filter set.
This technique typically results in incorporation of transgene DNA into one of the two daughter blastomeres of the first cell division [14]. Since the first cleavage separates prospective left and right sides of the embryo, half of each embryo inherits the transgene, and half does not. This accounts for the observed transgene expression in one or two (but never all three) palps, and in otolith or ocellus but not both, as these two cells derive from opposite sides of the embryo [16].

Embryos, in situ hybridisation and immunohistochemistry

Culturing of *C. intestinalis* embryos and RNA in situ hybridisation were carried out as described [17]. For immunohistochemistry, embryos were raised and fixed as for in situ hybridisation. They were then rehydrated in phosphate buffered saline with 0.2% Triton X100 (PBT), and blocked in 20% sheep serum in PBT (blocking solution) overnight at 4ºC. Embryos were then incubated at 4ºC overnight in pre-blocked (4ºC overnight, diluted 1:2000 in blocking solution) anti-Ci-βγ-crystallin antibody (rabbit polyclonal; raised against recombinant protein using standard procedures), followed by 6 one hour washes in PBT at room temperature and a second overnight incubation at 4ºC in preblocked (4ºC overnight in blocking solution) secondary antibody (Alexifluor 594 donkey anti-rabbit IgG, Molecular Probes). Embryos were then washed 6 times for 1 hour each, and viewed under a confocal microscope at 594 nm.

1. Turner, D.L., and Weintraub, H. (1994). Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. Genes Dev *8*, 1434-1447.
2. Kroll, K.L., and Amaya, E. (1996). Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. Development *122*, 3173-3183.
3. Sparrow, D.B., Latinkic, B., and Mohun, T.J. (2000). A simplified method of generating transgenic *Xenopus*. Nucleic Acids Res *28*, e12.
4. Nieuwkoop, P.D., and Faber, N. (1994). Normal table of *Xenopus laevis* (Daudin) (New York: Garland Publishing Inc.).
5. Bateman, O.A., Lubsen, N.H., and Slingsby, C. (2001). Association behaviour of human betaB1-crystallin and its truncated forms. Exp Eye Res *73*, 321-331.
6. Leslie, A.G.W. (1992). In Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, Volume 26.
7. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.
8. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr *61*, 458-464.
9. Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph *14*, 51-55, 29-32.
10. Morris, R.J., Perrakis, A., and Lamzin, V.S. (2003). ARP/wARP and automatic interpretation of protein electron density maps. Methods Enzymol *374*, 229-244.
11. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.
12. McRee, D.E. (1999). XtalView/Xfit - A versatile program for manipulating atomic coordinates and electron density. J Struct Biol *125*, 156-165.
13. Vagin, A., and Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. Acta Crystallogr D Biol Crystallogr *56 Pt 12*, 1622-1624.
14. Corbo, J.C., Levine, M., and Zeller, R.W. (1997). Characterization of a notochotd-specific enhancer from the *Brachyury* promoter region of the ascidian, *Ciona intestinalis*. Development *124*, 589-602.
15. Mita-Miyazawa, I., Ikegami, S., and Satoh, N. (1985). Histospecific acetylcholinesterase development in the presumptive muscle cells isolated from 16-cell-stage ascidian embryos with respect to the number of DNA replications. J Embryol Exp Morphol *87*, 1-12.
16. Darras, S., and Nishida, H. (2001). The BMP/CHORDIN antagonism controls sensory pigment cell specification and differentiation in the ascidian embryo. Dev Biol *236*, 271-288.
17. Boorman, C.J., and Shimeld, S.M. (2002). Pitx homeobox genes in *Ciona* and amphioxus show left-right asymmetry is a conserved chordate character and define the ascidian adenohypophysis. Evol Dev *4*, 354-365.

Table S1 X-ray Diffraction Statistics

Parameter	Value	
Space group	C2	
Unit Cell	a) 93.0Å b) 29.8Å c) 57.3Å $\alpha = \gamma = 90^{O}$, $\beta = 121.5^{O}$	
Wavelength	0.933 Å (ID14-2, ESRF)	
Temperature of data collection	100K	
Molecules in asu.	2	
Matthews Coefficient	2.0	
Solvent Content	37.0% (v/v)	
	All Data (40.0 – 1.50Å)	High Resolution Bin (1.58 – 1.50Å)
Number of Reflections	277325	9285
Number of Unique Reflections	21220	2879
Multiplicity	3.6	3.3
Completeness	93.0%	87.7%
R_{merge}	13.3%	30.9%
<I>/<Sig I>	2.6	1.8

Table S2 Refinement Statistics

Parameter	Value
Number of reflections working (test) set	18383 (2442)
R_{cryst} (R_{Free}) after rigid body	33.1% (33.3%)
Final R_{cryst} (R_{Free})	20.0% (23.9%)
Overall Figure of Merit	0.840
Number of atoms in final refinement round	2807
Rmsd bond lengths	0.007Å
Rmsd bond angles	1.13^{O}
Rmsd main chain B-factor bond (angle)	$1.65Å^2$ ($2.35Å^2$)
Rmsd side chain B-factor bond (angle)	$2.10Å^2$ ($3.19Å^2$)
Ramachandran plot: Most favoured region	87.2%
Ramachandran plot: Additionally allowed region	12.8%

Table S3 Comparison of calcium binding between ci-βγ-crystallin and spherulin 3a.

Ci-βγ-crystallin		Spherulin 3a	
First Motif			
Ligand	Distance (/Å) (Chain A/B)	Ligand	Distance (/Å) (Chain A/B)
Glu 7 O	2.3 / 2.3	Lys 19 O	2.3 / 2.3
Ile 33 O	2.4 / 2.4	Val 46 O	2.2 / 2.4
Ser 35 OG	2.4 / 2.3	Ser 48 OG	2.3 / 2.0
Asp 75 OD1	2.5 / 2.3	Asp 89 OD1	2.3 / 2.3
Glu 7 OE1 (symmetry)	2.2 / 2.3 B / A Glu 7		
HOH 2 / 5	2.4 / 2.5	HOH 37Z / 29Y	2.6 / 2.7
HOH 9 / 14	2.3 / 2.5	HOH 71Z / 50Y	2.5 / 2.2
Second Motif			
Asp 48 O	2.3 / 2.3	Lys 62 O	2.4 / 2.3
Glu 76 O	2.4 / 2.3	Ala 90 O	2.5 / 2.2
Ser 78 OG	2.3 / 2.3	Ser 92 OG	2.4 / 2.1
Asp 32 OD1	2.5 / 2.5	Asp 45 OD1	2.5 / 2.5
Glu 76 OE1	2.6 / 2.8	HOH 28 Z	2.8
HOH 52 / 27	3.6 / 2.3	HOH 36Z / 26 Y	2.6 / 2.6
		HOH 41Z / 30 Y	2.7 / 2.7

Ligands in equivalent positions are on the same line. Ligands are listed in the order main chain, side chain and finally water.

Figure S1. The single domain calcium bound Ci-βγ-crystallin in layers in the lattice.