Evolution of crystallins for a role in the vertebrate eye lens

Christine Slingsby,^{1*} Graeme J. Wistow,² and Alice R. Clark¹

¹Department of Biological Sciences, Crystallography, Institute of Structural and Molecular Biology, Birkbeck College, Malet Street, London WC1E 7HX, United Kingdom

²Section on Molecular Structure and Functional Genomics, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892-0608

Received 21 December 2012; Revised 16 January 2013; Accepted 17 January 2013

DOI: 10.1002/pro.2229

Published online 7 February 2013 proteinscience.org

Abstract: The camera eye lens of vertebrates is a classic example of the re-engineering of existing protein components to fashion a new device. The bulk of the lens is formed from proteins belonging to two superfamilies, the α -crystallins and the $\beta\gamma$ -crystallins. Tracing their ancestry may throw light on the origin of the optics of the lens. The α -crystallins belong to the ubiquitous small heat shock proteins family that plays a protective role in cellular homeostasis. They form enormous polydisperse oligomers that challenge modern biophysical methods to uncover the molecular basis of their assembly structure and chaperone-like protein binding function. It is argued that a molecular phenotype of a dynamic assembly suits a chaperone function as well as a structural role in the eye lens where the constraint of preventing protein condensation is paramount. The main cellular partners of α -crystallins, the β - and γ -crystallins, have largely been lost from the animal kingdom but the superfamily is hugely expanded in the vertebrate eye lens. Their structures show how a simple Greek key motif can evolve rapidly to form a complex array of monomers and oligomers. Apart from remaining transparent, a major role of the partnership of α -crystallins with β - and γ -crystallins in the lens is to form a refractive index gradient. Here, we show some of the structural and genetic features of these two protein superfamilies that enable the rapid creation of different assembly states, to match the rapidly changing optical needs among the various vertebrates.

Keywords: α -crystallin; β -crystallin; γ -crystallin; refractive index; small heat shock protein; chaperone; stress resistance; optics; protein–protein interactions; cataract

Abbreviations: ACD, α -crystallin domain; AIM1, absent in melanoma 1 protein; CryoEM, cryoelectron microscopy; EDSP, epidermal differentiation-specific protein; ETX, epsilon toxin; Frosty MAS NMR, freezing rotational diffusion of protein solutions at low temperature and high viscosity magic-angle-spinning nuclear magnetic resonance spectroscopy; SAX, small-angle X-ray scattering; sHsp, small heat shock protein.

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Christine Slingsby, Department of Biological Sciences, Crystallography, Institute of Structural and Molecular Biology, Birkbeck College, Malet Street, London WC1E 7HX, UK. E-mail: C.Slingsby@mail.cryst.bbk.ac.uk

Grant sponsor: Medical Research Council, UK; Grant number: G0801846; Grant sponsor: European Synchrotron Radiation Facility (Grenoble, France); Grant sponsor: Diamond Light Source (Oxfordshire, UK); Grant sponsor: Soleil, France (partially funded through BIOSTRUCT-X).

Diversity of Eyes

The tremendous evolutionary advantages conferred by the ability to respond to light are evident in the success of species from the unicellular, with simple eyespots, to vertebrates with image forming eyes.¹ In the animal kingdom, the six phyla (out of 35) that are most widespread and numerous are those that have image forming eyes while many others have light sensing systems. In all eyes, light is absorbed by related members of the opsin superfamily arrayed in either rhabdomeric or ciliary photoreceptor cells that transduce the optical signal through distinct mechanisms.² Beyond this basic level of light sensitivity, the structures and optics of eyes are extremely diverse. Eyes can be single or compound, gathering, and directing light onto the photoreceptors of the retina with pinholes, lenses, cylinders, or mirrors. Vertebrates use a camera eye with a cellular lens situated behind a curved cornea. In fish, underwater, the lens alone provides almost all the focusing power, while in terrestrial species, in air, the cornea provides most focusing power and the lens is mainly used for fine control of image formation.

The vertebrate lens is derived embryologically from an invaginated ectodermal epithelium, the lens vesicle, and grows throughout life by the orderly proliferation and differentiation of epithelial cells into layers of extremely elongated fiber cells.³ Cell organization is important for lens transparency and focusing, but most of the refractive power of the lens is conferred by high concentrations of proteins, with any highly abundant protein being designated a crystallin. The most widespread and apparently ancient crystallins found in vertebrate lineages are the α -, β -, and γ -crystallins. Nonchordates, even those with superficially similar cellular lenses, use quite different proteins. This shows that lenses arose independently, relatively late in evolution and means the crystallins must have been selected from proteins with pre-existing functions. In the case of α -crystallins the original function is very likely a role in protein homeostasis as they belong to the family of small heat shock (stress) proteins that are ubiquitous across all domains of life and most cellular types.⁴⁻⁶ The β - and γ -crystallins are not related to α -crystallins but are members of another protein superfamily of restricted phylogenetic and tissue distribution. In vertebrates, β - and γ -crystallins are highly expressed in the lens, with low levels found in some other eye tissues, particularly in different retinal cell types.⁷⁻¹⁰ In many vertebrate lineages, the optical properties of the lens have been also modified to adapt to environmental constraints by loss of some crystallins (generally γ -crystallins) and by independent recruitment of other proteins which, surprisingly, are usually well characterized enzymes.¹¹ Thus crystallins are all proteins that

have been adapted from their original function to be constituents of the optics of animal eyes.

The recruitment of a protein as an eye lens crystallin required altering gene regulation to achieve the high protein concentrations needed to bend light. This probably occurred through step-wise modification to promoters of genes, such as those for stress proteins or enzymes, which were already expressed at lower levels in the lens.¹² Not any protein will do. Crystallins need to pack closely enough and uniformly enough to eliminate spaces and concentration discontinuities on the scale of a light wavelength to create a transparent cytoplasmic medium. Refraction and transparency in the lens both require high protein solubility. This requires crystallins to maintain packing at high concentration while resisting crystallization and phase separation.^{13,14} Furthermore, since lens fiber cells lose all their organelles as they mature, presumably to reduce light scattering,¹⁵ crystallins need to be extremely long-lived. Indeed, those laid down in the embryo are retained in the central region of the lens throughout life while exposed to light. With no turnover and with limited available capacity to degrade or refold damaged or aggregated proteins, crystallins have evolved to be stable and to help stabilize each other. Crystallins have also coevolved with "beaded filaments" which are intermediate filaments unique to the eye lens.¹⁶

Recruitment of Small Heat Shock Proteins (sHsps) to the Vertebrate Eye Lens

sHsps can bind to denatured substrate proteins in vitro and prevent them from forming light scattering aggregates^{6,17}; as such they present obvious advantages for a lens. In eukaryotes, sHsp genes are upregulated by heat shock transcription factors allowing their expression to be upregulated by stress.¹⁸ This would appear to be a good starting point for the evolution of high levels of expression required for a lens crystallin. Furthermore, sHsps assemble to form large oligomers that are very polydisperse, a feature that lowers the risk of crystallization and phase separation. sHsps are found in bacteria, archaea, and eukaryotes, with the family members generally increasing in number in the more complex organisms. The two sHsp family members that are expressed as α -crystallins in the eye lens are generally the most closely related, with the gene duplication event likely to have occurred around the time that vertebrates arose. In humans the genes (*CRYAA* and *CRYAB*) encode proteins α A- and α B-crystallin (HSPB4 and HSPB5) that are 53% identical,¹⁹ and they coassemble in the lens to form α -crystallin of molecular weight around 800 kDa. Nanoelectrospray Mass Spectrometry shows α B-crystallin forms a wide range of oligomeric forms containing between 10 and 40 subunits.²⁰ *CRYAB* has retained the heat shock

element in the promoter region, and it is widely expressed in other tissues, particularly in other long lived, elongated cells like muscle and the neuroglia. *CRYAA* is more highly expressed in lens than *CRYAB*, but is not responsive to stress. It seems that α A-crystallin is the result of gene duplication and specialization for some aspects of the lens role, at the expense of the ancestral role as a sHsp.

Assembly of sHsps from the Alpha-Crystallin Domain (ACD) Dimer

All sHsp sequences have a single ACD of around 90 amino acids (Pfam PF00011: Hsp20/alpha-crystallin family), flanked by variable sequence extensions.²¹ The first sHsp crystal structure, HSP16.5 from the archaeal *Methanococcus jannaschii*, showed a 24-mer assembled into a hollow octahedron.²² This was followed by the structure of *Triticum aestivum* (wheat) HSP16.9, which assembled into two interlocking hexameric discs.²³ These regular oligomeric structures show how the ACD folds into a β -sandwich [Fig. 1(A)], similar but not identical in topology to an immunoglobulin domain, with the domain forming dimers [Fig. 1(B)] that are further assembled by the N- and C-terminal extensions into the dodecamer [Fig. 1(C)]. However, resolving struc-

tures from the polydisperse vertebrate assemblies is more problematical.

Crystal structures of dimers from single domain metazoan sequences were only resolved by removing sequence extensions,^{24–27} see Table I. The α B-crystallin ACD domain folds into a β -sandwich (comprising strands B2-B9) that is similar to the wheat monomer (Fig. 2). The mode of dimerization in non-metazoans, in which a B6 strand exchange occurs between paired domains, differs from that observed in metazoans where the B6 strand forms an antiparallel dimer interface (Fig. 2). The metazoan arrangement results in a deep groove at the interface [Fig. 3(A)]. A conserved sequence motif (I-X-I/V) in the C-terminal extension plays a conserved role in higher assembly, by binding into two pockets (one for each hydrophobic side chain) formed at the B4/B8 edges of the β -sandwich domain from a different dimer (Figs. 1 and 2). The different modes of dimerization place the pockets from the two dimers in a different spatial organization relative to each other (Fig. 2), influencing the geometry of the final assembly.

X-ray and solution NMR spectroscopy structures²⁸ have been solved for several vertebrate ACDs from α A, α B-crystallin and Hsp20 (HSPB6), Hsp27 (HSPB1) with and without the C-terminal

Figure 1. A dodecameric sHsp assembly is held together by sequence extensions from the ACD. A: The complete monomer fold of one chain from wheat HSP16.9 [PDB 1gme] is shown, with the beta-sandwich structure of the ACD shown in yellow (except for strands B8 and B8 that are colored dark blue), and the N-terminal region is colored green. The hydrophobic side chains of the C-terminal sequence motif, which is I-Q-I in this sequence, are appended in yellow space-fill. B: The dimer, showing one complete chain pairing with a chain lacking a resolved N-terminal extension, is viewed with the approximate dyad aligned vertically. Note that the I-Q-I motif is in two orientations due to alternative linker conformations. The pockets between B4 and B8 strands in each domain are positioned on one side of the dimer, with the exchanged B6 strand and N-terminal region on the other side. C: The full assembly, which has point group 32 symmetry, is made from six dimers arranged to form two interlocking discs. The oligomer two-fold aligns with a crystallographic two-fold, and not the local (pseudo) dimer two-fold axis. One dimer, colored and orientated as in B, is shown docking I-Q-I motifs into pockets inside the B4/B8 edge (blue) strands of an adjacent dimer in a disc (the other I-Q-I extension is making the equivalent interaction with a dimer from the disc below). Three N-terminal helical extensions further stabilize the assembly by interacting with three N-terminal helical extensions from dimers in the disc below. The image was created in Pymol.

Table I. A Selection of Crystallin Protein Structures and Their Homologs

Crystallin ^a	Assembly	Method	PDB	Species and location
<i>Hsp20/alpha-crystallin family</i>				
α -crystallin (A 1–173 and B 1–175)	~800 KDa polymer	UC	–	<i>Homo sapiens</i> lens
α B-crystallin assembly (1–175)	10–40 mer	NanoMS	–	<i>Homo sapiens</i> lens and muscle
α B ACD (67–157)	Dimer	X-ray	2wj7	<i>Homo sapiens</i>
α B R120G ACD (67–157)	Dimer	X-ray	2ylz	<i>Homo sapiens</i>
α B ACD + C-term (68–162)	Runaway dimer	X-ray	3l1g	<i>Homo sapiens</i>
α A ACD + C-term (62–163)	Runaway dimer	X-ray	3l1f	<i>Bos taurus</i>
α B-crystallin assembly (1–175)	~500 KDa	SS NMR	2klr	<i>Homo sapiens</i> lens and muscle
α B-crystallin assembly (1–175)	~500 KDa	CryoEM	2ygd	<i>Homo sapiens</i> lens and muscle
Hsp20 (HSPB6) (65–162)	Dimer	X-ray	2wj5	<i>Rattus norvegicus</i> muscle
Hsp27 (HSPB1) (90–171)	–	X-ray	3q9p	<i>Homo sapiens</i> muscle and neural
Tsp 36 (2–313) 2 ACDs per chain	Dimer	X-ray	2bol	<i>Taenia saginata</i> oncosphere
Hsp 16.9 (3–151 and 45–151)	Six dimers	X-ray	1gme	<i>Triticum aestivum</i>
Hsp 16.5 (33–147) archaea	24 mer	X-ray	1shs	<i>Methanococcus jannaschii</i>
Hsp 16.5 plus engineered insert	48 mer	X-ray	4eld	<i>Methanococcus jannaschii</i>
Hsp 14.0 (17–123) archaea	24 mer	X-ray	3vqk	<i>Sulfolobus tokodaii</i>
<i>β-crystallin-like domains</i>				
Truncated β B1-crystallin (42–251)	Dimer	X-ray	1oki	<i>Homo sapiens</i> lens
β B2-crystallin (2–205)	Dimer and tetramer	X-ray	2bb2 and 1ytq	<i>Bos taurus</i> and <i>Homo sapiens</i> lens
β B2-crystallin (15–103)	Dimer of N-domains	X-ray	1e7n	<i>Mus musculus</i>
β B3-crystallin (23–199)	–	X-ray	3qk3	<i>Homo sapiens</i> lens
β A4-crystallin (8–196)	–	X-ray	3lwk	<i>Homo sapiens</i> lens
γ B-crystallin (2–175)	2-Domain monomer	X-ray	1amm and 2jdf	<i>Bos taurus</i> and <i>Homo sapiens</i> lens
γ B-crystallin (88–175)	C-Domain monomer	X-ray	1dsl	<i>Bos Taurus</i>
γ B-crystallin (88–173)	Dimer of C-domains	X-ray	1gam	<i>Bos Taurus</i>
γ C-crystallin (2–174)	2-Domain monomer	X-ray	2vtu	<i>Mus musculus</i> lens
γ D-crystallin (2–174)	2-Domain monomer	X-ray	1hk0	<i>Homo sapiens</i> lens
γ E-crystallin (2–174)	2-Domain monomer	X-ray	1m8u and 1a5d	<i>Bos taurus</i> and <i>Rattus norvegicus</i> lens
γ S-crystallin (2–178)	2-Domain monomer	NMR	1zwm and 2a5m	<i>Mus musculus</i> lens
γ S-crystallin (91–177)	Dimer of C-domains	X-ray	1ha4	<i>Homo sapiens</i>
Urochordate β -crystallin (2–83)	1-Domain monomer	X-ray	2bv2	<i>Ciona intestinalis</i> otolith and palp
Aim1 g1 (1022–1118)	Domain-1 monomer	X-ray	3cw3	<i>Homo sapiens</i> epithelial cells
Aim1 g5 (1416–1495)	Domain-5 monomer	NMR	2dad	<i>Homo sapiens</i> epithelial cells
Slime mold Spherulin 3A (2–101)	1-Domain dimer	X-ray	1hdf	<i>Physarum polycephalum</i> encystment
Archaeal β -crystallin (37–120)	1-Domain monomer	X-ray	3hz2	<i>Methanosarcina acetivorans</i>
Bacterial Protein S (2–173)	2-Domain monomer	NMR	1prp	<i>Myxococcus xanthus</i> spore coat
Bacterial Protein S (2–89)	N-Domain monomer	X-ray	1nps	<i>Myxococcus xanthus</i> spore coat
Bacterial β -domain (119–206)	1-Domain monomer	X-ray	3i9h	<i>Clostridium beijerinckii</i>
Bacterial β -domain (501–590)	1-Domain monomer	X-ray	3hzb	<i>Flavobacterium johnsoniae</i>

^a Protein sequence numbers are based on the Uniprot entry which includes N-terminal Met. CryoEM, Cryo-electron microscopy; NanoMS, Nano-electrospray Mass Spectrometry; SS NMR, Solid State Nuclear Magnetic Resonance Spectroscopy; UC, Ultracentrifugation.

extension containing the **I-X-I/V** motif (Table I). All the dimer crystal structures show a deep groove at the antiparallel (AP) dimer interface in which small molecules from the mother liquor are sometimes bound [Fig. 3(A)]. There is variation though in the register of this interface and in the interface between pockets and C-terminal extensions. In those crystal structures that include the conserved sequence motif (**I-X-I/V**) in the C-terminal extension, the motif is bound in the B4/B8 pockets of an adjacent dimer. The direction of binding of this motif in the pockets is of interest as the sequence is palindromic (ERTIPITRE) in mammalian α B-crystallins leading to the suggestion that it could bind in either orientation. Paradoxically, the orientation is the same in 3D structures of α B-crystallin, the archaeal and plant assemblies, whereas in (bovine) α A-crys-

tallin it is in the reverse direction. These structural features likely favored the selection of sHSPs as lens proteins, as the **I-X-I/V** and AP interfaces promote oligomer polydispersity by creating assembly diversity.²⁵

A Conserved Arginine at the Dimer Interface

One of the most conserved residues in sHSPs across all domains of life is an arginine in the B7 strand, corresponding to R120 in human α B-crystallin. In the metazoan dimer, this places an arginine at each end of the AP interface groove [Fig. 3(A)]. The first two disease-associated mutations discovered in the human α -crystallins are to this arginine residue. Since then, many families with inherited neuromuscular disease have been shown to carry mutations at the equivalent residue in HSPB1, HSPB5, and

Figure 2. The ACD monomer is very similar in sHsps from all kingdoms of life, but the mode of dimerization differs. The beta sandwich monomer domain is similar in human α B-crystallin and wheat HSP16.9. Sheet strands B2, B3, B6 (absent in human), B9 are colored in yellow, and B5 and B7 are colored in tan. The B4 and B8 strands, colored dark blue, form the pockets into which the I-X-I/V motif (in ball and stick), insert. The main differences are in the shape and length of the chain connecting B5 with B7. The wheat dimer forms by B6 strand swapping between monomers, whereas in human α -crystallin the B7 strand merges with B6 strand and dimerizes by self-association about a two-fold axis to form an extended beta-sheet. Note the different spatial orientations of the B4/B8 docking stations for the I-X-I/V sequence motifs.

HSPB8 or in the case of aA-crystallin/HSPB4, have cataract.^{29,30} In human α B-crystallin the mutation R120G causes both cataract and a skeletal muscle myopathy. The conserved arginines form ion pairs with acidic side chains across the dimer interface: patients that carry a mutation to the acidic side chain partner suffer the same disease phenotype. The crystal structure of the ACD of the R120G mutant [PDB: 2ylz] showed that the open groove at the AP interface was closed [Fig. 3(B)]. We speculated that the extended β -sheet across the interface is dynamic allowing regulated access to the groove, and that this is dampened by the mutation.³¹ Furthermore, the N-terminal sequence extensions, which are extremely variable in sHsps, might be (partially) occupying this groove in the assembly in normal (unstressed) conditions. When the solid state NMR structure of the full assembly of human α B-crystallin was determined [PDB: 2klr], although the I-X-I/V motifs in the C-terminal extensions were bound in the pockets in the same orientation as the crystal structure, the N-terminal region was largely invisible.³² Furthermore, the extended anti-parallel β -sheet at the AP interface was in the shape of a hyperbolic paraboloid in contrast with the flat shape of the sheet in the crystal structures,³¹ and this obliterated the groove.

A clue as to how these metazoan sHsps might bind denatured polypeptides came from the structure of the ACD from rat Hsp20 (HSPB6).²⁴ The crystal lattice was held together by sequence extensions binding in the AP groove and pockets from partner dimers [Fig. 3(C)]. This led us to suggest

that the groove may act in concert with the pockets, to form general binding sites for hydrophobic regions from unfolded protein chains, which could wrap around and cross between domains, in a similar manner to the way the N-terminal region double-wrapped ACDs [Fig. 3(D)] in the tapeworm Tsp36 dimer structure.³³

The Challenge of Alpha-Crystallin Polydispersity

The visualization of crystallographic snap shots of a small number of monodisperse sHsp oligomeric assemblies distantly related to the α -crystallins, and of various subassembly units of the α -crystallins themselves (Table I), is largely in agreement with the solid-state NMR structure of the human α B-crystallin homo-oligomer, and solution NMR peptide-binding studies.³⁴ These observations established how assembly appears to be driven by the conserved sequence motif (I-X-I/V) of the C-terminal extension docking into the B4/B8 pockets of a partner dimer, while the long and variable N-terminal extension is assigned to disorder and assumed to be contributing to much of the conformational polydispersity of the system. However, Frosty MAS (freezing rotational diffusion of protein solutions at low temperature and high viscosity magic-angle-spinning) NMR spectroscopy studies challenge these rules as they indicate that the C-terminal peptide motif hovers rather than docks into the B4/B8 pockets, and that the monomer structure including the N-terminal region does not change conformation with temperature.^{35,36} The dynamics of the sHsp assemblies are likely key to their chaperone function. Nano-electrospray mass spectrometry demonstrates how a regular dodecamer, like the wheat HSP 16.9, transforms on warming to a highly heterogeneous polydisperse system³⁷ and the technology can also be used to determine stoichiometries of sHsps with model oligomeric substrates.³⁸ It is suggested that a conformation-induced increase in assembly states for sHSPs is at the heart of the substrate-binding mechanism by creating a wide range of potential substrate binding sites. This fits the idea that sHsps act as “passive” chaperones by holding onto a denatured substrate when ATP levels are compromised.⁶

Single particle image reconstruction by electron microscopy, although dependent on the level of particle homogeneity,³⁹ has measured particle size, symmetry and organization of several nonmetazoan sHsps. CryoEM and spin-labeling have revealed that HSP16.5 from thermostable *Methanococcus jannaschii*, in which 24 ACDs are assembled by docking of C-terminal extensions into pockets, forms a cubo-octahedral shell that encapsulates the N-terminal extension.⁴⁰ Early particle reconstructions from negative stain and cryoEM images of human α B-crystallin homo-oligomers did not reveal symmetry, but gave an indication of a hollow shell.⁴¹ More recently,

Figure 3. The peptide-binding regions of ACDs. A: Looking down the two-fold axis of human α B-crystallin ACD dimer with strands embedded in transparent space fill rendition. The sHsp superfamily conserved arginine (R120) in each monomer ion pairs with D107 across the interface. It is proposed that the groove space is a candidate substrate-binding site for unfolded proteins, along with the B4/B8 pockets (in dark blue) [PDB: 2wj7]. B: The R120G mutation, that causes cataract and myopathy, leads to conformational and interface ion-pair changes that result in groove closure [PDB: 2y1z]. C: In the crystal lattice of the ACD dimer of rat Hsp20, sequence extensions are bound in the groove (green) and pockets (blue) [PDB: 2wj5]. D: In the tapeworm Tsp36, the chain has two ACDs (gray ribbon) and hence two sets of B4/B8 pockets (blue ribbon) and forms a dimer (not shown). The pockets from ACD1 are filled by an I-F-P sequence motif from the N-terminal region of the partner chain (blue ball and stick) and the pockets from ACD2 are filled by an I-Q-P sequence from the linker between the two domains (gray ball and stick). The long helical N-terminal region from one chain (green ball and stick) is shown wrapping around both domains [PDB: 2bol].

negative stain EM has produced a map of human α B-crystallin, interpreted as a 24-mer with tetrahedral 23 symmetry, with tentative placement of a dimer in the asymmetric unit.⁴² This arrangement was further explored with SAX scattering and modeling data to suggest an arrangement built from rings of trimers of dimers with some similarities to the X-ray derived oligomer structures of more distantly related sHsps.⁴³ However, the latest cryoEM studies have led to a radically different interpretation [PDB: 2ygd], in which a building block is also a trimer of dimers but arranged in a porin-like structure.⁴⁴

Measurements of oligomer size and shape by nanoMS led to the proposal of a more flexible arrangement in which α B-crystallin assembly was based around interconverting polyhedrons with dimers occupying the edges.²⁰ Models of α B-crystallin based on cryoEM and nanoMS technologies must grapple with the challenge of size distribution in which an ensemble of oligomers centering around a 28-mer can exist in equilibrium with a range of oligomers of even and odd stoichiometries distributed among 10 to 40-mers. A dynamic interconverting native assembly is an attractive mechanism for

the generation of unlimited substrate binding regions, though the distribution of docked versus hovering states for the (I-X-I/V) assembly straps is currently in dispute. While a degree of fluctuation is necessary for sHsp “holdase” function, the pockets and grooves would have to be more permanently filled in the substrate bound state.

As protein oligomers that are highly soluble yet defy the rules of symmetry, sHSPs are evidently quite resistant to crystallization and this may fit them well for the role of crystallin. However, it remains an interesting question whether the ancestral chaperone role of α -crystallin is really central to its role in lens.³⁰ Although α -crystallins are capable of protecting other crystallins from aggregation, many of the other crystallins are themselves very stable and, under most conditions, are less likely to need a chaperone than α -crystallin itself. Indeed, the adaptive pressure that gave rise to α A-crystallin early in the evolution of the vertebrate lens seems to have sacrificed aspects of the chaperone role in favor of increased specialization as a crystallin. A very similar thing has also occurred in the zebrafish (*Danio rerio*) in which (in addition to α A-crystallin) there are two α B-crystallin-like genes, one of which

(α B2) is widely expressed and has chaperone function while the other (α B1) is more specialized for lens and has reduced chaperone function.⁴⁵ It may be that the diversity and flexibility of assembly originally evolved for the role as a “holdase” chaperone, was the key attribute responsible for the recruitment of sHsps as lens α -crystallins.

The Lens $\beta\gamma$ -Crystallin Superfamily

The oligomeric β -crystallins and monomeric γ -crystallins, which belong to a common $\beta\gamma$ -crystallin superfamily, constitute around 50% of the wet weight of the cytoplasmic lens proteins in mammals. They are all comprised of two similar $\beta\gamma$ -crystallin domains each formed from two repeat Greek key motifs of around 40 residues defined by a sequence fingerprint. All vertebrates, including the jawless lamprey, have multimember β - and γ -crystallin families, characterized at the gene level by the correspondence of a protein motif to an exon in β -crystallins, while in most γ -crystallins, an exon corresponds to a domain.^{3,46} In mammals, there are six β -crystallin genes, belonging to two groups whose members are around 50% identical in protein sequence, and there are clear orthologues of these genes in birds and teleost fishes. Most mammals have eight genes for γ -crystallins. Six of these (γ A-F) are encoded in a cluster, presumably the result of gene duplication, with 70–98% sequence identity. These are predominantly expressed in the developing lens and are particularly prone to evolutionary loss. Fish have no orthologues of these genes but have large numbers of genes for γ M-crystallins that are absent from mammals. Two other γ -crystallins have better conservation. γ S-crystallin is found in all vertebrates and is expressed at high levels in adult human lens.^{3,46,47} γ N-crystallin is an evolutionary intermediate between the β - and γ -crystallins, with orthologous genes in fish, reptiles, mammals, and birds. It has a gene structure in which the N-domain is encoded like a γ -crystallin (one exon for two motifs) while the C-domain is encoded like a β -crystallin (one exon per motif). Mouse lens expresses low levels of γ N, but in humans the gene appears to be non-functional.⁴⁷ γ -crystallins may have a particular importance in determining the optical properties of the lens. They are extremely abundant and diverse in the very high refractive index fish lens. Recently, a comprehensive survey of the amino acid composition of all proteins has found that γ -crystallins have unusually high contents of residues with higher refractive index increments (dn/dc) which may make a significant contribution to increased refractive power of the lens.⁴⁸

In mammals, the distribution of γ -crystallins changes along the optical axis of the lens, being most abundant in the densest core region (the lens nucleus) and lower in abundance in more hydrated

cortical regions. Thus, the contribution of γ -crystallins parallels the gradient of refractive index that is required to reduce spherical aberration in the lens. Interestingly, γ A, γ E, γ F-crystallins in the lens nucleus exhibit phase separation at close to ambient temperatures.^{49,50} This tendency to self associate would seem to be a risk for lens transparency, but has presumably evolved to promote the very close packing of proteins in the densest regions of lenses. Indeed, while these proteins are expressed significantly in the hard, high refractive index lenses of mouse and rat, in humans (with more flexible, accommodating lenses) the *CRYGA* gene is expressed at low levels while *CRYGEP* and *CRYGFP* are pseudogenes. In contrast, γ S-crystallin is abundant in the softer, more hydrated, outer region of the lens and is the single most abundant γ -crystallin in humans.^{51,52} Deletion of the gene for γ S-crystallin in mice is associated with abnormal organization of F-actin throughout the lens cortex while, in vitro, mouse γ S-crystallin can stabilize F-actin.⁵³ This hints at a role for γ -crystallins in helping to maintain and stabilize oligomeric structures like actin and perhaps also other crystallins.

Mammalian γ -crystallins are thermodynamically extremely stable and resistant to photodamage.⁵⁴ Their surface polar residues are extensively involved in intramolecular ion pairs and hydrogen bonds. Interestingly, the γ A-F-crystallins have distinctively high ratios of Arg/Lys residues. The surface arginines seem to be key for solubility as evidenced by their propensity to cause congenital cataract when mutated.^{30,46} In the case of human γ D-crystallin, a surface mutation R37S created a lattice contact [PDB: 2g98] and led to crystals forming in the lens.⁵⁵ Paradoxically, a different point mutation in the same protein, R59H, led to loss of a lattice contact [PDB: 1h4a] compared with the wild-type in the same unit cell [PDB: 1hk0] and prickle shaped crystals in the lens.⁵⁶ A proline residue in the first Greek key motif of human γ D-crystallin also appears to be a hot spot for cataract related mutations since several families around the world carry dominant P24T mutations.⁵⁷ NMR spectroscopy [PDB: 2kfb] showed that the mutation had an impact on local conformational dynamics.⁵⁸

Structural Diversity of $\beta\gamma$ -Crystallins from a Common Architecture

The 3D structures of β - and γ -crystallins reveal monomeric and various oligomeric structures all elaborated from the highly conserved motif/domain foundation of the superfamily. They are all comprised of two similar $\beta\gamma$ -crystallin domains. Each domain is formed from two Greek key motifs characterized by an unusual folded β -hairpin that intercalate to form a β -sandwich with each pair of motifs arranged about an approximate dyad (Figs. 4

Figure 4. γ -crystallin structure recapitulates its ancestry from repeating motifs. Four Greek key motifs form two similar domains organized about a pseudo twofold axis. The first motif of each domain (A-type) is shown in dark blue and the second motif (B-type) is in cyan. B-type motifs provide the conserved tyrosine corner residues (shown in ball and stick) and tryptophan corners, as well as hydrophobic residues that form the “paired domain” interface. A-type motifs are more surface facing.

and 6). The second of each pair of motifs in $\beta\gamma$ -crystallin domains has characteristic tyrosine and tryptophan corner residues that underpin the β -sandwich fold. In the monomeric γ -crystallins, the two domains are connected by a short, bent linker polypeptide, and in the various crystal lattices from which they were solved, the domains are also paired about an approximate dyad (Fig. 4). This arrangement likely recapitulates the ancestry of these proteins, starting from formation of a homodimer of proto-motifs; after gene duplication this was succeeded by a more stable heterodimer of motifs solidified by gene fusion to form a domain; duplication and dimerization of domain heterodimers was eventually solidified by gene fusion giving rise to modern two domain proteins.^{59,60} Crystal structures of engineered single domains of β - and γ -crystallins recapitulate the presumed pairing (dimerization) of an ancestral single domain (Table I). All the structures have the second Greek key motif (B) of each domain at the interface between domains. As a result, motifs 2 and 4 in the complete chain are more conserved than A-type motifs 1 and 3, probably because they provide the hydrophobic residues at the core of the paired domain interface, as well as the corner tyrosine and tryptophan residues (Figs. 4 and 6). This also allows the more exposed motifs 1 and 3 to adapt to enhance appropriate interactions in the lens cytoplasm.

β -crystallins can be subdivided into two groups, the basic (β B1, β B2, and β B3) and acidic (β A3/1, β A2, and β A4). Like the γ -crystallins, they have two similar domains connected by a linker, but the β -crystallins all have N-terminal sequence extensions, some quite lengthy, while the basic ones also have C-terminal extensions.³ When expressed in vitro they assemble into homodimers and higher

Figure 5. Conserved interfaces in $\beta\gamma$ -crystallins. γ -crystallins are monomeric because the “paired domain” interface is intrachain. The same interface is present in all the resolved β -crystallins. However, in β B2-crystallin the linker between domains is extended so this interface forms between two different chains and thus creates a domain swapped dimer. In β B1, β B3, and β A4-crystallin, the subunit domains are paired as in γ -crystallins, and so these dimers need an additional interface. This new interface is the same as that observed in the β B2 tetramer formed in the crystal lattice (see Supporting Information Fig. 1). In the case of β B3-crystallin, the calculated assembly unit is a trimer, with the conserved dimer interface being formed in the crystal lattice (see Supporting Information Fig. 1).

order oligomers, although in the lens they preferentially occur as hetero-oligomers of great size and charge heterogeneity, and the larger in vivo assemblies cannot be reconstituted in vitro.⁶¹ The highest order assemblies always contain β B1-crystallin, which has the longest N-terminal extension of all. β B2-crystallin is generally the predominant β -crystallin and forms dimers and hetero-oligomers with all the β -crystallin chains.

When the X-ray structure of bovine β B2-crystallin was solved, it was striking to see that dimerization was effected by domain swapping.⁶² The pseudosymmetric domain pairing of two domains in γ -crystallins was faithfully recapitulated except that

Figure 6. Tracing the evolution of the $\beta\gamma$ -domain. An A-type Greek key motif (dark blue) followed by a B-type motif (cyan) fold to form the basic $\beta\gamma$ -domain in all lens β - and γ -crystallins. The fold is characterized by tryptophan and tyrosine corners, appended in ball and stick, and there are no calcium binding sites. The B-type motif has hydrophobic surface residues to drive the domain pairing seen in Figures 4 and 5. In the urochordate and archaea domains, both corners are present, the domain pairing interface is absent, and calcium binding sites are loaded (yellow spheres). In bacteria, the motifs are permuted, the tryptophan corner is absent, the loaded calcium binding sites are present, and the domain pairing interface is absent.

the “paired domain” interaction was intermolecular and the connecting peptide was extended (Fig. 5). This domain swapping was robust to sequence extension truncation,⁶³ unfolding-refolding [PDB: 1ytq]⁶⁴ but not to domain permutation [PDB: 1bdz].⁶⁵ The determinants for domain swapping of β B2-crystallin are hard wired into the sequence of the domains, and are not dependent on the linker or extensions.

In contrast, the homo-oligomeric structure for sequence truncated human β B1-crystallin showed intrachain pairing of domains as in the γ -crystallins, with the monomer–monomer (dimer) interface recapitulating the domain swapped dimer–dimer interface found in the crystal lattices of bovine and human β B2-crystallin (Fig. 5 and Supporting Information Fig. S1).⁶⁶

Coordinates have been deposited for homo-oligomers of β B3 and β A4-crystallins from genomics consortia, and they again show intrachain domain pairing like γ -crystallin and β B1-crystallin (Fig. 5). In the human β B3-crystallin structure [PDB: 3qk3], the calculated biological assembly is a new contact form: a trimer. However, exploration of the lattice also reveals a dimer contact similar to that in β B1-crystallin (Fig. 5 and Supporting Information Fig. S1). In human β A4-crystallin [PDB: 3lwk], the calculated biological assembly has the same dimer arrangement as β B1-crystallin (Fig. 5). Only rudimentary stumps of the sequence extensions can be seen in these β -crystallin structures.

In general, it seems that there is a conserved interface in β -crystallins. β B2-crystallin stands out for its strong preference for domain swapping. It is noticeable that β B2-crystallin is thermodynamically the least stable of all crystallins, is involved in dynamic subunit exchange with other β -crystallins *in vitro*⁶¹ and can act as a “chaperone” to stabilize/coas-

semble other β -crystallins *in cellulo*.⁶⁷ Indeed, homo-oligomers may not exist in the lens: what interfaces exist in the hetero-oligomers are not known.

Overall, it seems that the double Greek key $\beta\gamma$ domain has duplicated and rapidly expanded the family of 2-domain proteins whilst creating a wide variety of assemblies with a limited number of interfaces. Like the α -crystallins, the important constraints are packing at high protein concentration combined with polydispersity to avoid phase separation, crystallization, or precipitation and the maintenance of short range order on the scale of light wavelength over a lifetime.

The Urochordate and Cephalochordate $\beta\gamma$ -Crystallins

$\beta\gamma$ -crystallins in the lens date back to the roots of the vertebrate lineage. Clues to their origins may lie in close relatives of vertebrates in the chordate sub-phyla that diverged before the development of the vertebrate camera eye. A single domain $\beta\gamma$ -crystallin is present in the urochordate sea squirt, *Ciona intestinalis*, called ciona-crystallin.⁶⁸ The organization of the gene, including the phases of the exon/intron junctions, is identical to that of the first half of vertebrate β -crystallin genes, with two exons each encoding one Greek key motif. The structure of the domain shows the typical pseudosymmetric fold of paired Greek key motifs, but unlike lens $\beta\gamma$ -crystallins, two calcium atoms are bound by the domain (Fig. 6). A sequence fingerprint **D/N-X-X-S** in each motif is associated with calcium binding with each motif fingerprint contributing one conserved side chain to form two half binding sites. Ciona-crystallin is found in two locations in the swimming larval form of the sea squirt. It is synthesized and retained inside anterior palp cells that secrete adhesives for sticking the larva onto surfaces ready

for metamorphosis into the sessile state. These urochordates also have a pair of anterior pigmented sister cells in the larval head, the otolith and the ocellus that function in guidance by responding to depth and light, respectively.⁶⁹ Ciona-crystallin is primarily located in the otolith,⁶⁸ which has significant levels of calcium.⁷⁰ The ocellus is a ciliary-based photoreceptor system homologous to the vertebrate retina. Further support for an ancestral relationship came from the remarkable observation that ciona-crystallin gene promoter was able to target reporter gene expression to the vertebrate visual system.⁶⁸

Vertebrate $\beta\gamma$ -crystallin domain pairing depends on a conserved pattern of hydrophobic residues in the second and fourth motifs (Fig. 4). This set of residues is not conserved in ciona-crystallin, consistent with its monomeric nature in solution.⁶⁸ Analysis of the genome of cephalochordate *Branchiostomata floridae* (amphioxus), which is considered to be ancestral to the vertebrates and urochordates,⁷¹ shows it has genes for “primitive” looking $\beta\gamma$ -crystallins that share protein features of both urochordates and vertebrates.⁷² There are several genes encoding two domain proteins, and they have the typical tyrosine and tryptophan corner arrangement contributed from residues in motifs 2 and 4 for each domain (Fig. 6). They lack the domain pairing hydrophobics in motif 4, but two sequences have them in motif 2, leading to the possibility that these N-terminal domains might pair with each other (Table I). Two of the proteins have calcium-binding fingerprints in both motifs 3 and 4 making it likely that the C-terminal domains in these proteins bind calcium. Surprisingly, the cephalochordate $\beta\gamma$ -crystallin genes do not show the motif/exon mapping seen in vertebrates and urochordates. The widespread occurrence of this mapping indicates, it might be ancestral and that intron loss has occurred in some lineages. Amphioxus does not have surface eyes, but there are several photosensitive organs cells belonging to both the rhabdomic and ciliary systems, and recent molecular and cellular evidence supports homology of the amphioxus frontal eye with the vertebrate eye.⁷³ It will be interesting to determine if, when and where the amphioxus $\beta\gamma$ -crystallins are expressed.

Nonlens Vertebrate $\beta\gamma$ -Crystallins

Another way of addressing the question of the original function of $\beta\gamma$ -crystallins is to look at the structure and function of relatives expressed outside the lens. A 12-motif, six domain sequence, with exon encoded motifs like β -crystallin and ciona crystallin genes is present as part of a much larger protein called “absent in melanoma 1” (AIM1), which is expressed in several epithelial cell types.⁷⁴ 3D structures have been solved for the first⁷⁵ and fifth domains (Table I). The $\beta\gamma$ -crystallin domains have all

the characteristics of the lens $\beta\gamma$ -crystallins, and a model of their assembly has been proposed.⁷⁴ The domains display embellishments to the protein fold, and they do not have the calcium binding sites. The AIM1 gene is present in all vertebrates, including the hagfish *P. Marinus*, but has not been found in urochordate or cephalochordate genomes.⁷² AIM1 contains a long, repetitive N-terminal region in addition to a C-terminal ricin B-type trefoil lectin domain. The function is not known, but deletion of the AIM1 gene region is associated with loss of typical fibroblast-like cell structure and increased malignancy in melanoma cells.⁷⁴ This hints at a role in control of cell architecture.

Other relatives have been found in amphibians. The protein Ep37/EDSP in the newt *Cynops* has a pair of domains with typical vertebrate $\beta\gamma$ -crystallin sequence and is associated with plasma membrane in developing larval epidermis.^{76,77} The N-terminal domain of Ep37/EDSP also has the sequence fingerprint for calcium binding, so far the only sighting of such a domain in vertebrates. C-terminal to the $\beta\gamma$ -crystallin domains is a Clostridium epsilon toxin ETX domain. A single $\beta\gamma$ -crystallin domain followed by an ETX domain (B2BRT1) is also found in the skin of a giant fire-bellied toad, called $\beta\gamma$ -CAT, and is a toxic skin secretion.⁷⁸ This protein forms a dimer with a trefoil domain protein, suggestive of a simpler version of the arrangement in AIM1.

Microbial $\beta\gamma$ -Crystallin-Like Sequences

Early analyses spotted the four-fold signature sequence fingerprint of lens $\beta\gamma$ -crystallin in microbial spore coat protein sequences.^{60,79} Crystal structures of one of the two domains from a bacterium, *Myxococcus Xanthus* and a single domain, obligate dimer from the mycetezoan *Physarum polycephalum* showed they have very similar calcium-binding sites^{80,81} that are also very similar to ciona-crystallin.⁶⁸ In contrast with the vertebrate, urochordata and cephalochordate $\beta\gamma$ -crystallin domains, the bacterial sequence has a tyrosine corner in the first, not the second motif, and there is no tryptophan corner (Fig. 6). Similar single domain $\beta\gamma$ -crystallin-like sequences are now known to be more widespread in the bacterial world, embedded in much larger sequences and several calcium-bound crystal structures have been solved⁸² (Table I). However, in archaea, single domain calcium-binding $\beta\gamma$ -crystallin-like sequences are found with the motifs and corners arranged as in vertebrates⁸² (Fig. 6). Synthesis of several of the calcium loaded $\beta\gamma$ -crystallin-like microbial proteins is associated with survival under conditions of duress. Similarly, sHsps are abundantly expressed in life-cycle stages of mycobacteria,⁸³ parasitic worms,⁸⁴ and arthropods,⁸⁵ where they provide stress resistance, often in association with dehydration.

In addition to chordate and related species, there are genes in the poriferan demospongiae *Geodia cydonium* (Uniprot O18426),⁸⁶ and in the cnidarian sea anemone, *Nematostella vectensis* (XP_001633217), that encode $\beta\gamma$ -crystallin motifs and are predicted to fold into at least one $\beta\gamma$ -domain, based on the presence of the sequence fingerprints for the folded hairpin, and presence of tyrosine and tryptophan corner residues in the second motif. Half a D/N-X-X-S fingerprint is present in each Greek key motif, so there could also be one complete calcium binding site within the domain.

Do proteins in all these species share a common evolutionary origin? The observation of a calcium-bound single $\beta\gamma$ -crystallin-like domain with a vertebrate arrangement of motifs and corners in a member of the archaea,⁸² is at least consistent with a common origin for $\beta\gamma$ -crystallin motifs or domains in animals. If the poriferan, cnidarian, and vertebrate sequences are ancestrally related, the lack of relatives in other major animal phyla such as Nematoda, Arthropoda, Mollusca, Platyhelminthes, Enteropneusta, and Echinodermata, suggests that the genes for $\beta\gamma$ -domains were repeatedly lost. Indeed, the same thing has happened for other proteins of mammalian lens. The lens is characterized by high levels of an ancient member of the glutamine synthetase superfamily, lensin, which is specifically expressed in terminally differentiating lens fiber cells.⁸⁷ It is expressed in the lens of species from fish to birds but is absent from arthropods and other phyla, yet it has several relatives in the genome of the sea urchin, an echinoderm.

Conclusions

Lenses act to gather and focus light. Some simple organisms have patches of photoreceptors that react rapidly to shadows of predators but gather no other information whereas a lens (or another optical solution) gives directional information and in concert with more advanced brain functions can give the eye the ability to form information rich images.¹ Since the fish cornea, in direct contact with the optically dense medium of the water, provides little refractive power, the evolution of the camera eye in fishes involved acquiring a dense, spherical lens with a steep gradient of refractive index to provide focusing power whilst minimizing spherical aberration.⁸⁸ Many vertebrates have multifocal lenses to compensate for chromatic aberration.⁸⁹

To achieve the advantages of the high protein concentrations necessary for refraction, evolutionary processes selected for proteins that could easily be expressed in the newly developing tissue, that could resist phase separation, aggregation, and crystallization and that were highly stable. In the vertebrate lens, the core groups of proteins that arose were the α -crystallins, with the ability to form a wide and

dynamic range of structures built from just two polypeptides; and the β - and γ -crystallins that generated diversity rapidly by repeated duplication and conservation of interfaces.

α -crystallins are expressed throughout the lens and are highly polydisperse. This, rather than their ancestral chaperone-like function may be the major reason for their recruitment as crystallins. Interestingly, although α -crystallins are stable, soluble proteins, they make an increasing contribution to the insoluble fraction of the mature transparent lens.⁹⁰ sHsp assemblies have been described based on polyhedra architecture⁹¹ and larger 3D nets could form by modulating the swapping angle of the I-X-I/V extensions. This was observed in switching between regular polyhedra in an engineered [PDB: 4eld] archaeal sHsp.⁹² α -crystallins may also grow in size by a self-chaperone-like protein binding function. This change in the size distribution of α -crystallin along the light path may be part of a mechanism to adjust the refractive index gradient of the ever-growing eye lens.³⁰

Multiple β -crystallins are able to form a wide range of hetero-oligomers of different sizes. As their expression levels are regulated throughout lens development and throughout life, they too can modulate the packing density of the lens cytoplasm with different oligomer forms to adjust the gradient of refractive index. Monomeric γ -crystallins are also developmentally regulated with different contents in different regions of the lens. Their surfaces are highly specialized, with networks of polar interactions. Indeed, it has been suggested that different γ -crystallins maintain characteristic molecular dipoles that may allow them to adopt specific orientations relative to other lens components, perhaps modulating their interactions and packing: changes that disrupt these dipoles may be associated with cataract.⁹³

Over time, as species in different lineages moved into new environments (such as the emergence onto land) and had different requirements for lens refractive power and flexibility, the composition of the lens adjusted. Several times γ -crystallins were reduced in expression or lost and other available proteins, such as the taxon-specific enzyme crystallins, were recruited.¹¹ Indeed, this occurred in the human lineage with substantial loss of γ -crystallins and the embryonic expression of the enzyme betaine-homocysteine methyltransferase as ψ -crystallin.^{46,94} However, the interplay of α -crystallins and $\beta\gamma$ -crystallins is the basis of the optical properties of the vertebrate lens.

References

1. Land MF, Nilsson D-E (2002) Animal eyes. Oxford: Oxford University press.
2. Arendt D, Tessmar-Raible K, Snyman H, Dorresteyn AW, Wittbrodt J (2004) Ciliary photoreceptors with a

- vertebrate-type opsin in an invertebrate brain. *Science* 306:869–871.
3. 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.
4. Haslbeck M, Franzmann T, Weinfurter D, Buchner J (2005) Some like it hot: the structure and function of small heat-shock proteins. *Nat Struct Mol Biol* 12: 842–846.
5. Mchaourab HS, Godar JA, Stewart PL (2009) Structure and mechanism of protein stability sensors: Chaperone activity of small heat shock proteins. *Biochemistry* 48: 3828–3837.
6. Basha E, O'Neill H, Vierling E (2012) Small heat shock proteins and alpha-crystallins: dynamic proteins with flexible functions. *Trends Biochem Sci* 37:106–117.
7. Sinha D, Esumi N, Jaworski C, Kozak CA, Pierce E, Wistow G (1998) Cloning and mapping the mouse Crygs gene and non-lens expression of gammaS-crystallin. *Mol Vis* 4:8.
8. Organisciak D, Darrow R, Gu X, Barsalou L, Crabb JW (2006) Genetic, age and light mediated effects on crystallin protein expression in the retina. *Photochem Photobiol* 82:1088–1096.
9. Andley U P (2007) Crystallins in the eye: function and pathology. *Prog Retin Eye Res* 26:78–98.
10. Parthasarathy G, Ma B, Zhang C, Gongora C, Zigler JS, Duncan MK, Sinha D (2011) Expression of beta A3/A1-crystallin in the developing and adult rat eye. *J Mol Histol* 42:59–69.
11. Wistow G (1993) Lens crystallins-gene recruitment and evolutionary dynamism. *Trends Biochem Sci* 18: 301–306.
12. Cvekl A, Yang Y, Chauhan BK, Cveklova K (2004) Regulation of gene expression by Pax6 in ocular cells: a case of tissue-preferred expression of crystallins in lens. *Int J Dev Biol* 48:829–844.
13. Delaye M, Tardieu A (1983) Short-range order of crystallin proteins accounts for eye lens transparency. *Nature* 302:415–417.
14. Benedek GB (1997) Cataract as a protein condensation disease. The Proctor Lecture. *Invest Ophthalmol Vis Sci* 38:1911–1921.
15. Bassnett S, Shi Y, Vrensen GFJM (2011) Biological glass: structural determinants of eye lens transparency. *Philos Trans R Soc Lond B Biol Sci* 366:1250–1264.
16. Song SH, Landsbury A, Dahm R, Liu YZ, Zhang QJ, Quinlan RA (2009) Functions of the intermediate filament cytoskeleton in the eye lens. *J Clin Invest* 119: 1837–1848.
17. Horwitz J (1992) Alpha-crystallin can function as a molecular chaperone. *Proc Natl Acad Sci USA* 89: 10449–10453.
18. de Thonel A, Le Mouel A, Mezger V (2012) Transcriptional regulation of small HSP-HSF1 and beyond. *Int J Biochem Cell Biol* 44:1593–1612.
19. Kappé G, Franck E, Verschuure P, Boelens WC, Leunissen JAM, de Jong WW (2003) The human genome encodes 10 alpha-crystallin-related small heat shock proteins: HspB1-10. *Cell Stress Chaperones* 8:53–61.
20. Baldwin AJ, Lioe H, Hilton GR, Baker LA, Rubinstein JL, Kay LE, Benesch JLP (2011) The polydispersity of α B-crystallin is rationalized by an interconverting polyhedral architecture. *Structure* 19:1855–1863.
21. Kriehuber T, Rattei T, Weinmaier T, Bepperling A, Haslbeck M, Buchner J (2010) Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *FASEB J* 24:3633–3642.
22. Kim KK, Kim R, Kim SH (1998) Crystal structure of a small heat-shock protein. *Nature* 394:595–599.
23. van Montfort RLM, Basha E, Friedrich KL, Slingsby C, Vierling E (2001) Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat Struct Biol* 8: 1025–1030.
24. Bagnris C, Bateman OA, Naylor CE, Cronin N, Boelens WC, Keep NH, Slingsby C (2009) Crystal structures of alpha-crystallin domain dimers of alpha B-crystallin and Hsp20. *J Mol Biol* 392:1242–1252.
25. Laganowsky A, Benesch JLP, Landau M, Ding LL, Sawaya MR, Cascio D, Huang QL, Robinson CV, Horwitz J, Eisenberg D (2010) Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci* 19:1031–1043.
26. Laganowsky A, Eisenberg D (2010) Non-3D domain swapped crystal structure of truncated zebrafish alphaA crystallin. *Protein Sci* 19:1978–1984.
27. Baranova EV, Weeks SD, Bukach OV, Gusev NB, Strelkov SV (2011) Three dimensional structure of alpha-crystallin domain dimers of human small heat shock proteins HSPB1 and HSPB6. *J Mol Biol* 411:110–122.
28. Jehle S, van Rossum B, Stout JR, Noguchi SM, Falber K, Rehbein K, Oschkinat H, Kleivit RE, Rajagopal P (2009) Alpha B-crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. *J Mol Biol* 385:1481–1497.
29. Sacconi S, Feasson L, Antoine JC, Pecheux C, Bernard R, Cobo AM, Casarin A, Salviati L, Desnuelle C, Urtizbera A (2012) A novel CRYAB mutation resulting in multisystemic disease. *Neuromuscul Disord* 22:66–72.
30. Clark AR, Lubsen, NH, Slingsby C (2012) sHSP in the eye lens: crystallin mutations, cataract and proteostasis. *Int J Biochem Cell Biol* 44:1687–1697.
31. Clark AR, Naylor CE, Bagnris C, Keep NH, Slingsby C (2011) Crystal structure of R120G disease mutant of human alphaB-crystallin domain dimer shows closure of a groove. *J Mol Biol* 408:118–134.
32. Jehle S, Rajagopal P, Bardiaux B, Markovic S, Kuhne R, Stout JR, Higman VA, Kleivit RE, van Rossum BJ, Oschkinat H (2010) Solid-state NMR and SAXS studies provide a structural basis for the activation of alpha B-crystallin oligomers. *Nat Struct Mol Biol* 17: 1037–1042.
33. Stamler R, Kapp G, Boelens W, Slingsby C (2005) Wrapping the alpha-crystallin domain fold in a chaperone assembly. *J Mol Biol* 353:68–79.
34. Delbecq SP, Jehle S, Kleivit R (2012) Binding determinants of the small heat shock protein, α B-crystallin: recognition of the 'I χ I' motif. *Embo Journal* 31: 4587–4594.
35. Baldwin AJ, Hilton GR, Lioe H, Bagnris C, Benesch JLP, Kay LE (2011) Quaternary dynamics of alphaB-crystallin as a direct consequence of localised tertiary fluctuations in the C-Terminus. *J Mol Biol* 413:310–320.
36. Baldwin AJ, Walsh P, Hansen DF, Hilton GR, Benesch JLP, Sharpe S, Kay LE (2012) Probing dynamic conformations of the high-molecular-weight alpha B-crystallin heat shock protein ensemble by NMR spectroscopy. *J Am Chem Soc* 134:15343–15350.
37. Stengel F, Baldwin AJ, Painter AJ, Jaya N, Basha E, Kay LE, Vierling E, Robinson CV, Benesch JLP (2010) Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proc Natl Acad Sci USA* 107:2007–2012.
38. Stengel F, Baldwin AJ, Bush MF, Hilton GR, Lioe H, Basha E, Jaya N, Vierling E, Benesch JLP (2012) Dissecting heterogeneous molecular chaperone complexes

- using a mass spectrum deconvolution approach. *Chem Biol* 19:599–607.
39. Orlova EV, Saibil HR (2011) Structural analysis of macromolecular assemblies by electron microscopy. *Chem Rev* 111:7710–7748.
 40. Koteiche HA, Chiu S, Majdorch RL, Stewart PL, McHaourab HS (2005) Atomic models by cryo-EM and site directed spin labeling: application to the N-terminal region of Hsp16.5. *Structure* 13:1165–1171.
 41. Haley DA, Horwitz J, Stewart PL (1998) The small heat-shock protein, alpha B-crystallin, has a variable quaternary structure. *J Mol Biol* 277:27–35.
 42. Peschek J, Braun N, Franzmann TM, Georgalis Y, Haslbeck M, Weinkauff S, Buchner J (2009) The eye lens chaperone alpha-crystallin forms defined globular assemblies. *Proc Natl Acad Sci USA* 106:13272–13277.
 43. Jehle S, Vollmar BS, Bardiaux B, Dove KK, Rajagopal, P, Gonen T, Oschkinat H, Klevit R. (2011) N-terminal domain of alphaB-crystallin provides a conformational switch for multimerization and structural heterogeneity. *Proc Natl Acad Sci USA* 108:6409–6414.
 44. Braun N, Zacharias M, Peschek J, Kastenmüller A, Zou J, Hanzlika M, Haslbeck M, Rappsilber J, Buchner J, Weinkauff S (2011) Multiple molecular architectures of the eye lens chaperone alphaB-crystallin elucidated by a triple hybrid approach. *Proc Natl Acad Sci USA* 108:20491–20496.
 45. Smith AA, Wyatt K, Vacha J, Vihtelic TS, Zigler JS, Wistow GJ, Posner M (2006) Gene duplication and separation of functions in alpha B-crystallin from zebrafish (*Danio rerio*). *FEBS J* 273:481–490.
 46. Wistow G (2012) The human gene families. *Hum Genomics* 6:26.
 47. Wistow G, Wyatt K, David L, Gao C, Bateman O, Bernstein S, Tomarev S, Segovia L, Slingsby C, Vihtelic T (2005) γ N-crystallin and the evolution of the $\beta\gamma$ -crystallin superfamily in vertebrates. *FEBS J* 272:2276–2291.
 48. Zhao HY, Brown PH, Magone MT, Schuck P (2011) The molecular refractive function of lens gamma-crystallins. *J Mol Biol* 411:680–699.
 49. Broide ML, Berland C R, Pande J, Ogun OO, Benedek G B (1991) Binary- liquid phase separation of lens protein solutions. *Proc Natl Acad Sci USA* 88:5660–5664.
 50. Norledge BV, Hay RE, Bateman OA, Slingsby C, Driessen HPC (1997) Towards a molecular understanding of phase separation in the lens: a comparison of the X-ray structures of two high T_c γ -crystallins, γE and γF , with two low T_c γ -crystallins, γB and γD . *Exp Eye Res* 65:609–630.
 51. Lampi KJ, Ma Z, Shih M, Shearer TR, Smith JB, Smith DL, David LL (1997) Sequence analysis of $\beta A3$, $\beta B3$, and $\beta A4$ crystallins completes the identification of the major proteins in young human lens. *J Biol Chem* 272:2268–2275.
 52. Wistow G, Bernstein SL, Wyatt MK, Behal A, Touchman JW, Bouffard G, Smith D, Peterson K (2002) Expressed sequence tag analysis of adult human lens for the NEI-Bank project: over 2000 non-redundant transcripts, novel genes and splice variants. *Mol Vis* 8:161–163.
 53. Fan J, Dong L, Mishra S, Chen YW, Fitzgerald P, Wistow G (2012) A role for gamma S-crystallin in the organization of actin and fiber cell maturation in the mouse lens. *FEBS J* 279:2892–2904.
 54. Chen J, Callis PR, King J (2009) Mechanism of the very efficient quenching of tryptophan fluorescence in human gamma D- and gamma S-crystallins: the gamma-crystallin fold may have evolved to protect tryptophan residues from ultraviolet photodamage. *Biochemistry* 48:3708–3716.
 55. Knoch S, Brynda J, Asfaw B, Bezouška K, Novák P, Rezáčová P, Ondrová L, Filipec M, Sedláček J, Elleder M (2000) Link between a novel human γD -crystallin allele and a unique cataract phenotype explained by protein crystallography. *Hum Mol Genet* 9:1779–1786.
 56. Basak A, Bateman O, Slingsby C, Pande A, Asherie N, Ogun O, Benedek GB, Pande J (2003) High-resolution x-ray crystal structures of human γD crystallin (1.25Å) and the R58H mutant (1.15Å) associated with aculeiform cataract. *J Mol Biol* 328:1137–1147.
 57. Evans P, Wyatt K, Wistow GJ, Bateman OA, Wallace BA, Slingsby C (2004) The P23T cataract mutation causes loss of solubility of folded γD -crystallin. *J Mol Biol* 343:435–444.
 58. Jung J, Byeon IJ, Wang Y, King J, Gronenborn AM. (2009) The structure of the cataract-causing P23T mutant of human gammaD-crystallin exhibits distinctive local conformational and dynamic changes. *Biochemistry* 48:2597–609.
 59. Blundell TL, Lindley PF, Miller L, Moss DS, Slingsby C, Tickle IJ, Turnell WG, Wistow GJ (1981) The molecular structure and stability of the eye lens: X-ray analysis of γ -crystallin II. *Nature* 289:771–777.
 60. Wistow G (1990) Evolution of a protein superfamily-relationships between vertebrate lens crystallins and microorganism dormancy proteins. *J Mol Evol* 30:140–145.
 61. Bateman OA, Sarra R, van Genesen ST, Kappé G, Lubsen NH, Slingsby C. (2003) The stability of human acidic β -crystallin oligomers and hetero-oligomers. *Exp Eye Res* 77:409–422.
 62. 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.
 63. Norledge BV, Trinkl S, Jaenicke R, Slingsby C (1997) The X-ray structure of a mutant eye lens $\beta B2$ -crystallin with truncated sequence extensions. *Protein Sci* 6:1612–1620.
 64. Smith MA, Bateman OA, Jaenicke R, Slingsby C (2007) Mutation of interfaces in domain swapped human $\beta B2$ -crystallin. *Protein Sci* 16:615–625.
 65. Wright G, Basak AK, Wieligmann K, Mayr E-M, Slingsby C (1998) Circular permutation of $\beta B2$ -crystallin changes the hierarchy of domain assembly. *Protein Sci* 7:1280–1285.
 66. van Montfort RLM, Bateman OA, Lubsen NH, Slingsby C (2003) Crystal structure of truncated human $\beta B1$ -crystallin. *Protein Sci* 12:2606–2612.
 67. Marin-Vinader L, Onnekink C, van Genesen ST, Slingsby C, Lubsen NH (2006) In vivo heteromer formation. Expression of soluble beta A4-crystallin requires coexpression of a heteromeric partner. *FEBS J* 273:3172–3182.
 68. 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.
 69. Horie T, Sakurai D, Ohtsuki H, Terakita A, Shichida Y, Usukura J, Kusakabe T, Tsuda M (2008) Pigmented and nonpigmented ocelli in the brain vesicle of the ascidian larva. *J Comp Neur* 509:88–102.
 70. Sakurai D, Goda M, Kohmura Y, Horie T, Iwamoto H, Ohtsuki H, Tsuda I (2004) The role of pigment cells in the brain of ascidian larva. *J Comp Neurol* 475:70–82.
 71. Delsuc F, Brinkmann H, Chourrout D, Philippe H (2006) Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439:965–968.
 72. Kappé G, Purkiss AG, van Genesen ST, Slingsby C, Lubsen NH (2010) Explosive expansion of gamma-

- crystallin genes in the ancestral vertebrate. *J Mol Evol* 71:219–230.
73. Vopalensky P, Pergner J, Liegertova M, Benito-Gutierrez E, Arendt D, Kozmik Z (2012). Molecular analysis of the amphioxus frontal eye unravels the evolutionary origin of the retina and pigment cells of the vertebrate eye. *Proc Natl Acad Sci USA* 109:15383–15388.
 74. Ray ME, Wistow G, Su Y A, Meltzer PS, Trent J M (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.
 75. 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.
 76. Takabatake T, Takahashi TC, Takeshima K (1992) Cloning of an epidermis-specific cynops cDNA from neurula library. *Dev Growth Differ* 34:277–283.
 77. Wistow G, Jaworski C, Vasantha Rao P (1995) A non-lens member of the $\beta\gamma$ -crystallin superfamily in a vertebrate, the amphibian cynops. *Exp Eye Res* 61:637–639.
 78. Liu S-B, He Y-Y, Zhang Y, Lee W-H, Qian J-Q, Lai R, Jin Y (2008) A novel non-lens $\beta\gamma$ -crystallin and trefoil factor complex from amphibian skin and its functional implications. *PLoS One* 3:e1770.
 79. 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.
 80. Wenk M, Baumgartner R, Holak TA, Huber R, Jaenicke R, Mayr EM (1999) The domains of Protein S from *Myxococcus xanthus*: structure, stability and interactions. *J Mol Biol* 286:1533–1545.
 81. Clout NJ, Kretschmar M, Jaenicke R, Slingsby C (2001) Crystal structure of spherulin S3a homodimer in the calcium-loaded form sheds light on the evolution of the eye lens $\beta\gamma$ -crystallin Greek-key domain fold. *Structure* 9:115–124.
 82. Aravind P, Mishra A, Suman SK, Jobby MK, Sankaranarayanan R, Sharma Y (2009) The beta gamma-crystallin superfamily contains a universal motif for binding calcium. *Biochemistry* 48:12180–12190.
 83. Sherman DR, Voskuil M, Schnappinger D, Liao RL, Harrell MI, Schoolnik GK (2001) Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha-crystallin. *Proc Natl Acad Sci USA* 98: 7534–7539.
 84. Moxon JV, LaCourse EJ, Wright HA, Perally S, Prescott MC, Gillard JL, Barrett J, Hamilton JV, Brophy PM (2010) Proteomic analysis of embryonic *Fasciola hepatica*: characterization and antigenic potential of a developmentally regulated heat shock protein. *Vet Parasitol* 169:62–75.
 85. King AM, MacRae TH (2012) The small heat shock protein p26 aids development of encysting *artemia* embryos, prevents spontaneous diapause termination and protects against Stress. *Plos One* 7 Issue: 8: e43723.
 86. Krasko A, Müller IM, Müller WEG (1997) Evolutionary relationships of the metazoan $\beta\gamma$ -crystallins, including that from the marine sponge *Geodia cydonium*. *Proc Roy Soc B* 264:1077–1084.
 87. Wyatt K, White HE, Wang LC, Bateman OA, Slingsby C, Orlova EV, Wistow G (2006). Lengsin is a survivor of an ancient family of class I glutamine synthetases re-engineered by evolution for a role in the vertebrate lens. *Structure* 14:1823–1834.
 88. Pierscionek BK, Regini JW (2012) The gradient index lens of the eye: An opto-biological synchrony. *Prog Ret Eye Res* 31:332–349.
 89. Gustafsson OSE, Collin SP, Kroger RHH (2008) Early evolution of multifocal optics for well-focused colour vision in vertebrates. *J Exp Biol* 211:1559–1564.
 90. Ortwerth BJ, Sharma KK, Olesen PR (1992) The effect of urea on the aggregate state and elastase inhibitor activity of the water-insoluble fraction from bovine and human lens. *Exp Eye Res* 54: 573–581.
 91. Hilton GR, Lioe H, Stengel F, Baldwin AJ, Benesch JLP (2012) Small heat-shock proteins: paramedics of the cell. *Topics in current chemistry*. Berlin Heidelberg: Springer-Verlag.
 92. Mchaourab HS, Lin YL, Spiller BW (2012) Crystal structure of an activated variant of small heat shock protein HSP16.5. *Biochemistry* 51:5105–5112.
 93. Purkiss AG, Bateman OA, Wyatt K, Wilmarth WA, David LL, Wistow GJ, Slingsby C (2007) Biophysical properties of γ C-crystallin in human and mouse eye lens: the role of molecular dipoles. *J Mol Biol* 372: 205–222.
 94. Rao PV, Garrow TA, John F, Garland D, Millian NS, Zigler JS (1998) Betaine-homocysteine methyltransferase is a developmentally regulated enzyme crystallin in rhesus monkey lens. *J Biol Chem* 273:30669–30674.
 95. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797.