

# Structure of the crystallins

CHRISTINE SLINGSBY, NAOMI J. CLOUT

## Abstract

The lens is formed from two protein superfamilies, the  $\alpha$ - and  $\beta\gamma$ -crystallins. Representative three-dimensional structures show they both have a basic 2- $\beta$ -sheet domain fold, with the  $\beta\gamma$ -domain being made from two intercalating Greek keys. X-ray structures of monomeric  $\gamma$ -crystallins and simple oligomeric  $\beta$ -crystallins show how multiple gene duplications can give rise to highly symmetrical assemblies based on paired domains. These protein folds have been engineered by directed mutagenesis to investigate the roles of the critical region in domain pairing and assembly. Inherited human cataracts have been described that are associated with representatives of each of the crystallin protein families. Mutations to certain  $\beta$ - and  $\gamma$ -crystallin genes cause expression of truncated polypeptides that would not be expected to fold properly; instead they would randomly aggregate causing light scattering. As crystallin proteins are not renewed, age-related cataract is a gradual accumulation of small changes to pre-existing normal proteins. The precise sites of post-translational modifications are now being mapped to the various crystallins.

*Key words* Cataract, Crystallin, Domain fold, Evolution, Eye lens, Protein structure

The molecular complexity underlying vision is encapsulated within well-ordered and distinctive morphological features. Now that mapping and sequencing the human genome is well under way, there will be rapid advances in understanding how such specialised systems function when all the protein components are defined. It is already known that almost every gene preferentially expressed in photoreceptors has mutations associated with retinal cell degeneration,<sup>1</sup> thus giving ophthalmology a crucial role in gene discovery. It is possible that mutations in any of the approximately 14 crystallin genes expressed in the lens could result in childhood cataract. However, it is the special feature of the lens whereby crystallins are made and then retained throughout life that makes cataract mainly a problem of ageing of many normal, non-renewable proteins.<sup>2</sup>

Protein-protein interactions are at the heart of the biophysical properties of crystallins as it is their concentration that provides both the refractive power of the lens and its transparency.<sup>3,4</sup> The crystallins, enclosed within the well-organised lens fibre cells, have to survive the life-span of the organism and so must be stable. Yet with age they lose solubility when there is little protein machinery around for their disposal. Loss of solubility and stability of the  $\alpha$ ,  $\beta$  and  $\gamma$ -crystallins by any means can lead to light scattering and cataract.<sup>5</sup>

The  $\alpha$ -crystallins are very large oligomers assembled from around 40 copies of two types (A and B) of sequence-related 20 kDa subunits.<sup>6</sup> The C-terminal 100 residues of these subunits are related to the small heat-shock proteins and examples of this ' $\alpha$ -crystallin domain fold' are found throughout the kingdoms of life. The  $\beta$ -crystallins are also oligomeric but more complex, in that they can consist of up to seven sequence-related polypeptides ranging in size from approximately 20 to 30 kDa. The  $\gamma$ -crystallins are a family of 21 kDa monomeric proteins and they share around 30% sequence identity with the  $\beta$ -crystallins. Representative three-dimensional crystal structures are available for all these classes of proteins.<sup>5,7</sup>

## The $\beta\gamma$ -crystallin domain fold

All the  $\beta$ - and  $\gamma$ -crystallin genes encode polypeptides that fold into two similar domains and each domain folds into two similar Greek key motifs. In  $\beta$ -crystallin genes the motifs are encoded in separate exons while in  $\gamma$ -crystallins exons code for whole domains.<sup>8</sup> Each Greek key motif consists of around 40 amino acid residues that fold into four consecutive anti-parallel  $\beta$ -strands, known as **a**, **b**, **c** and **d** strands (Figs. 1, 2). Two consecutive Greek key motifs are intercalated in a symmetrical way about an approximate dyad to form a wedge-shaped  $\beta$ -sheet sandwich filled with hydrophobic side chains (Fig. 2). The amino acid sequence encodes critical basic architectural features that allow this domain to fold. For example, the tertiary domain fold is stabilised by two folded hairpin loops (formed from residues between **a** and **b** strands in each motif) that require glycine and serine residues at critical positions of the fold (Figs. 1, 2). Each domain has one tyrosine corner that is present only in the second motif

C. Slingsby ✉  
N.J. Clout  
Birkbeck College  
Department of  
Crystallography  
Laboratory of Molecular  
Biology  
Malet Street  
London WC1E 7HX, UK

The financial support of the Medical Research Council is gratefully acknowledged

hB2 N-terminal extension ASDHQTQAGKPKQSLN  
 bB2 N-terminal extension ASDHQTQAGKPKQPLN

#### N-terminal domain

	← a <sub>1</sub> →	← b <sub>1</sub> →	← c <sub>1</sub> →	← d <sub>1</sub> →	← a <sub>2</sub> →	← b <sub>2</sub> →	← c <sub>2</sub> →	← d <sub>2</sub> →
hB2	PKIIIFEQENFQ <b>GH</b> SHELN-GPCPNLKE-TGV-EKAG <b>S</b> VLVQA		GPWVGYEQAN <b>CKGE</b> QFVFEKGE <b>Y</b> PRWDSWTSSRRRTD <b>SL</b> SLRPIK					
bB2	PKIIIFEQENFQ <b>GH</b> SHELN-GPCPNLKE-TGV-EKAG <b>S</b> VLVQA		GPWVGYEQAN <b>CKGE</b> QFVFEKGE <b>Y</b> PRWDSWTSSRRRTD <b>SL</b> SLRPIK					
gGB	GKITFYEDRGFQ <b>GH</b> CHYECS-SDCPNLQP-Y-F-SRC <b>S</b> IRVDS		GCWMLYERPNYQ <b>GH</b> QYFLRRGD <b>Y</b> PDYQQWGM-F-NDSIR <b>S</b> CRLLIP					
	1	20	40	60	80			

#### connecting peptide

hB2 VDSQE  
 bB2 VDSQE  
 bGB QHTGT

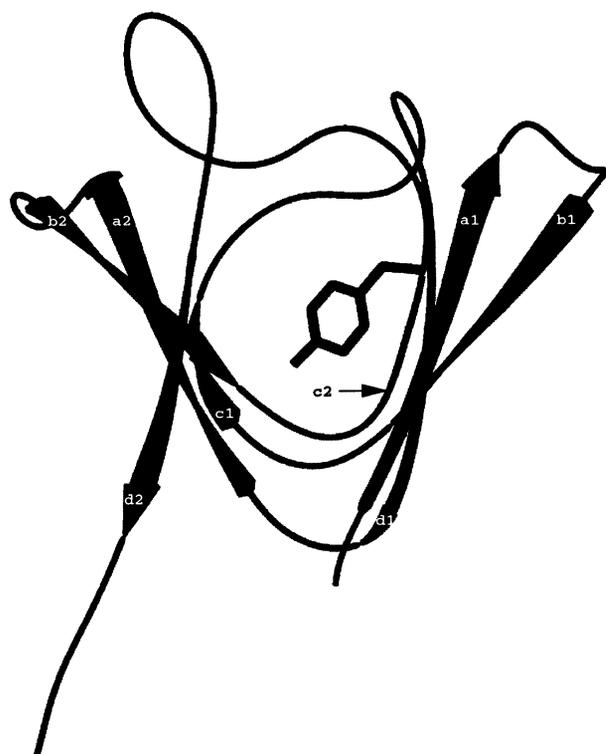
#### C-terminal domain

	←	←	←	←	←	←	←	←
hB2	HKIILYENPNFT <b>G</b> KKMEIIDDVPSFHA-HGYQEKV <b>S</b> SVRVQS		GTWVGYYQPGYR <b>GLQ</b> YLLEKGD <b>Y</b> KDSSDFGA-P-HPQVQ <b>S</b> VRRIR					
bB2	HKITLYENPNFT <b>G</b> KKMEVIDDDVPSFHA-HGYQEKV <b>S</b> SVRVQS		GTWVGYYQPGYR <b>GLQ</b> YLLEKGD <b>Y</b> KDSGDFGA-P-QPQVQ <b>S</b> VRRIR					
gGB	FRMRIYERDDFR <b>Q</b> MSEIT-DDCPSLQDRFHL-TEV <b>H</b> SLNVLE		GSWVLYEMPSYR <b>GRQ</b> YLLRPGE <b>Y</b> RRYLDWGA-M-NAKVG <b>S</b> LRRVM					
	100	120	140	160				

#### C-terminal extension

hB2 DMQWHQRGAFHPSN  
 bB2 DMQWHQRGAFHPSS  
 gGB DFY  
 174

**Fig. 1.** Sequence alignment of  $\beta$ B2- and  $\gamma$ B-crystallin modules (extensions, domains, connecting peptides/linker). The N- and C-terminal domain sequences of human (h) and bovine (b)  $\beta$ B2- and  $\gamma$ B-crystallins are also aligned vertically. The division between Greek key motifs is indicated by |. The numbering is based on the  $\gamma$ B sequence. The  $\beta$ -sheet strands are indicated by horizontal arrows and are labelled in bold lower-case lettering for motifs 1 and 2. Conserved residues mentioned in the text are highlighted in bold. Truncation sites are indicated by the symbol  $\perp$  below the equivalent residues of the  $\gamma$ B sequence.



**Fig. 2.** The  $\beta\gamma$ -domain domain fold. A schematic diagram of the fold of the polypeptide backbone of the N-terminal domain of bovine  $\beta$ B2 showing how the  $\beta$ -strands (thick arrows) of Greek key motifs 1 and 2 are intercalated to form a wedge-shaped domain from a pair of stacked  $\beta$ -sheets, drawn using the program SETOR.<sup>51</sup> The loop between the a and b strands forms a folded  $\beta$ -hairpin structure and is dependent on the conservation of a glycine (within the hairpin) and a serine (on the d strand) in each motif. Tyrosine 62 from the tyrosine corner of motif 2 has been appended.

(Figs. 1, 2). Point mutations to these highly conserved residues that maintain this fold are likely to destabilise the domain structure or prevent it from folding at all. An unfolded domain will aggregate randomly to form light scattering centres.

#### The monomeric $\gamma$ -crystallins

The three-dimensional structures of  $\gamma$ -crystallins reveal the next level of assembly of the single domain fold. In  $\gamma$ -crystallins the N- and C-terminal domains from the same polypeptide chain associate with each other in a symmetrical way. Domain pairing is driven by a patch of hydrophobic residues on the surface of each isolated domain and these get buried in the interface between domains (Fig. 3). The symmetry is important because it means that both hydrophobic patches are equally buried. The specific orientation of the domains is facilitated by a set of surrounding polar interactions, that includes conserved glutamines 54 and 143 in motifs in 2 and 4 (Figs. 1, 3).

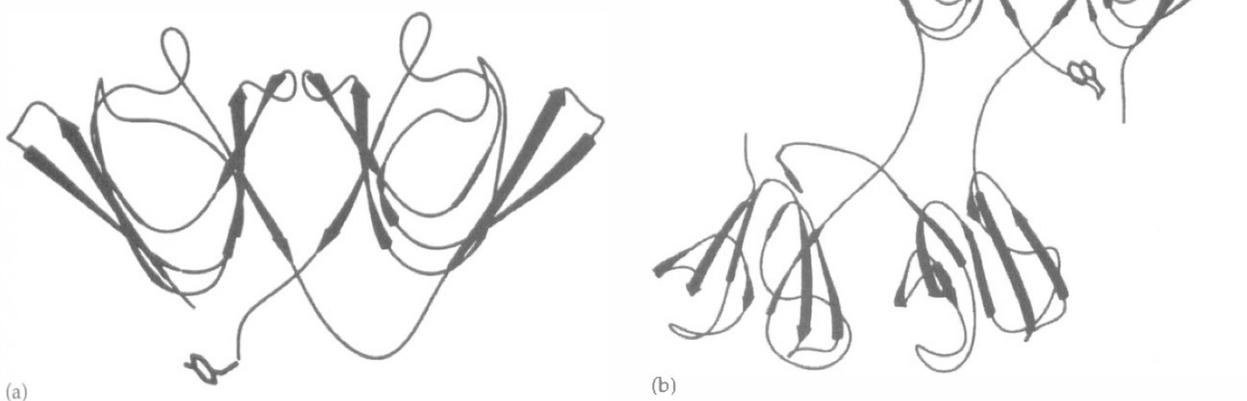
Protein engineering experiments have been performed to probe some of the sequence determinants of this two-domain assembly. When the X-ray structure of the first  $\gamma$ -crystallin was determined, it was recognised that the protein most probably evolved from a gene coding for a single domain that subsequently duplicated and fused, as was also suggested by the presence of an intron in the linker region. Protein-engineered single domains do not associate into a dimer of domains (like  $\gamma$ -crystallin) at low protein concentration, showing that



**Fig. 3.** The  $\gamma$ -crystallins have two domains. The domains are covalently linked by the connecting peptide between motifs 2 and 3. The two domains are held together non-covalently in this orientation by hydrophobic residues from the  $\beta$ -sheet strands from motifs 2 and 4, shown appended in black, and these are surrounded by polar residues, shown appended in grey. The topologically equivalent glutamines 54 and 143 are shown at the top of the interface region.

the linker enforces a local concentration that favours higher assembly.<sup>9</sup> When the single C-terminal domain structure of  $\gamma$ B-crystallin was solved a remarkable finding was the sensitivity of the symmetrical assembly to the C-terminal extension.<sup>10</sup> All  $\gamma$ -crystallins have a short C-terminal extension that crosses back to the N-terminal domain (Fig. 4a). When the engineered  $\gamma$ B C-terminal domain retains the intact C-terminal extension, then the single domain molecule is barely soluble. The crystal lattice structure shows that higher-domain assembly is now driven by interactions that bury the hydrophobic C-terminal extension in a 'head to tail' manner that prohibits symmetrical 'head to head'

domain pairing of the hydrophobic domain patch. Removal of just the one C-terminal tyrosine residue (Fig. 4a) restores proper pairing.<sup>10</sup> This dramatic demonstration of how solubility is related to higher-domain assembly is made more pertinent by the observation that  $\gamma$ -crystallin domains have been found in human cataractous lenses that are cleaved in the linker region.<sup>11</sup>



**Fig. 4.** The  $\beta$ - and  $\gamma$ -crystallins have evolved by domain swapping. (a) The two-domain  $\gamma$ -crystallin molecule is viewed perpendicular to the approximate twofold axis relating the N-terminal (on the left) and C-terminal domains. The covalent linker is kinked to allow the domains to pair. The C-terminal tyrosine residue is appended. (b) The N- and C-terminal domains of a  $\beta$ B2 polypeptide are separated from each other as the linker peptide is extended. The dimer is formed by the N-terminal domain of one subunit interacting with the C-terminal domain from a second subunit, while the C-terminal domain of the first subunit interacts with the N-terminal domain of the second subunit. The two sets of paired domains in the  $\beta$ B2 dimer have identical interfaces, and use the topologically equivalent residues to  $\gamma$ -crystallins (Fig. 3). Domain swapping thus creates a four-domain assembly, equivalent to two  $\gamma$ -crystallin molecules. Only a few residues of the N- and C-terminal extensions are visualised in the crystal structure, the remainder are disordered in water channels in the crystal.

The family of  $\gamma$ -crystallins can be divided into three groups. Expression of early genes  $\gamma$ E and  $\gamma$ F results in the corresponding proteins forming the bulk of the fetal nuclear region of the adult lens.  $\gamma$ B and  $\gamma$ C genes are switched on for longer and the corresponding proteins occupy a more intermediate region of the lens, while  $\gamma$ S is only expressed after birth and so this protein is restricted to the cortical region. The human  $\gamma$ E and  $\gamma$ F genes harbour early stop codons<sup>12</sup> and are fortunately not expressed, as the corresponding severely truncated  $\gamma$ -crystallin polypeptides would not be able to fold. Solubility studies have shown that the early proteins readily phase separate, causing cold cataract.<sup>13</sup> It is thought that this behaviour reflects their facility for tight packing and so is a structural feature associated with forming high refractive index gradients. Analyses of several X-ray structures of  $\gamma$ -crystallins indicated that arginine-containing ion-pairs played a role in this behaviour.<sup>14</sup> The reliance of humans on accommodation has necessitated a softer lens and this may explain why our early  $\gamma$ -genes have been effectively switched off.<sup>15</sup>  $\gamma$ S-crystallin does not phase separate and this may be due to the much lower levels of arginine.

There are subtle sequence differences between  $\gamma$ S and the other  $\gamma$ -crystallins in the domain interface region that may be reflected in the lower thermodynamic stability or altered biophysical properties of  $\gamma$ S-crystallin. Non-covalent domain pairing is found in the crystal lattice structures of an engineered single N-terminal domain of  $\gamma$ B-crystallin<sup>16</sup> and an engineered C-terminal domain of  $\gamma$ S-crystallin,<sup>17</sup> but in both cases the symmetry was different from that seen in native  $\beta$  $\gamma$ -crystallin structures. Conversely, when point mutations were made to a key hydrophobic residue at the interface in two-domain  $\gamma$ B-crystallin, then the symmetry of domain pairing was preserved, although the stability of the whole molecule was reduced.<sup>18</sup>

### The oligomeric $\beta$ -crystallins

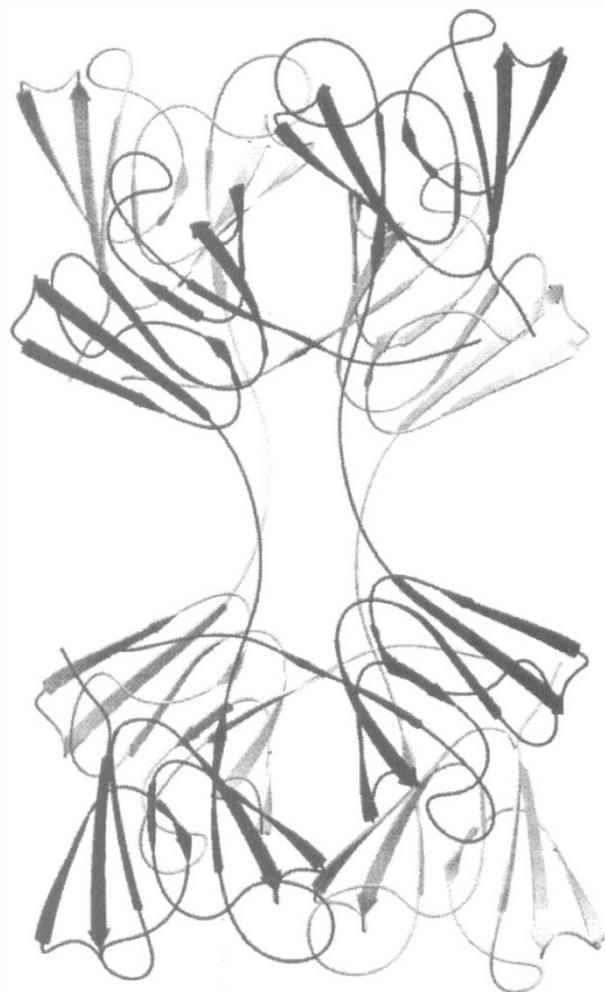
#### *$\beta$ -crystallin dimers and tetramers*

There are six  $\beta$ -crystallin genes and they differ from  $\gamma$ -crystallins in having long N-terminal sequence extensions.<sup>15</sup> They have been divided into groups that are called basic (B) and acidic (A), although the key sequence differences are based on group-specific domain insertions and the possession of long C-terminal extensions in basic  $\beta$ -crystallins.  $\beta$ -Crystallins have very similar domain structures to  $\gamma$ -crystallins (Fig. 2). The X-ray structure of  $\beta$ B2-crystallin showed how at least the basic branch of this family formed oligomers.<sup>19</sup> The trick was to use the same domain interface as the  $\gamma$ -crystallins, but to exploit the symmetry so that a two-domain assembly was changed into a four-domain assembly, by changing the conformation of the linker and allowing domains to swap partners between two molecules (Fig. 4b). Protein engineering experiments are trying to decipher the exact sequence rules that specify intramolecular versus intermolecular domain pairing.<sup>20–22</sup> It is to be noted that all  $\beta$ -crystallin

sequences have a proline at position 80. The X-ray structure of an engineered mutant  $\beta$ B2 in which both the N- and C-terminal extensions were truncated did not affect domain swapping and dimer formation in  $\beta$ B2.<sup>23</sup>

Although  $\beta$ B2-crystallin is a dimer in solution, in the highly concentrated environment of the crystal lattice the dimers associate into symmetrical tetramers (Fig. 5). A new face on the subunits of the dimer is buried in the new interface inside the tetramer.

The X-ray structure of  $\beta$ B2 dimer was intriguing in that the last observed residue in the X-ray structure at the C-terminal extension was close to the first observed residue at the N-terminus (Fig. 4b). This suggested that if these sequence termini were joined and the linker were cut, then a molecule like  $\gamma$ -crystallin might be formed. In other words, circularly permuting the domains, equivalent to domain/exon shuffling at the genetic level, might swap the domains back to intramolecular pairing. When such a molecule was created by protein engineering, the molecule remained a dimer in solution.<sup>24</sup> However, when the X-ray structure was solved,  $\gamma$ -crystallin-like domain pairing was observed,<sup>25</sup> and two such molecules formed a four-domain assembly identical to the top half of a  $\beta$ B2 tetramer (Fig. 5). When protein engineering creates new assemblies by simply changing the connections of folded domains, it is



**Fig. 5.** The  $\beta$ B2 tetramer is a dimer of  $\beta$ B2 dimers.

tempting to suggest that they also occur naturally. This has led to the hypothesis that acidic  $\beta$ -crystallin homodimers may in fact resemble the circularly permuted  $\beta$ B2 structure that evolved by exon shuffling.

### Higher oligomers

The several  $\beta$ -crystallin polypeptides in the cortical region of calf and adult bovine lens can be separated on gel filtration into three size classes, of which the smallest size is a dimer. Although purified  $\beta$ B2,  $\beta$ B3 or  $\beta$ A3 have been shown to be homodimers in dilute solution, in the lens fractions they mainly occur as heterodimers and higher oligomers.<sup>26</sup> In both human and bovine lenses, the bulk of the  $\beta$ -crystallin polypeptides are full-length only in young lenses. Truncation of  $\beta$ B1 and  $\beta$ A3/A1 begins early in life.<sup>27,28</sup> It has long been known that almost full-length  $\beta$ B1, the longest  $\beta$ -crystallin, is essential for assembly of the  $\beta$ H-crystallins. More recent data have shown that severely truncated  $\beta$ B1 polypeptides are present in the lower-molecular-weight  $\beta$ -crystallin oligomers, indicating that it is the N-terminal extension of  $\beta$ B1 that is key for the highest assembly.<sup>29</sup> Thus the lens can regulate the size distribution of the  $\beta$ -crystallin oligomers by either differential gene regulation of  $\beta$ B1 synthesis or by its post-translational truncation. The X-ray structures of  $\beta$ B2, solved in three different crystal lattices, show that its higher assembly is based on a tetramer structure. Hypothetical models have been proposed for  $\beta$ H-crystallin based on dimer–tetramer or tetramer–tetramer interactions.<sup>5</sup> One of the  $\beta$ B2 crystal lattice structures showed the N-terminal extension of  $\beta$ B2 mediating tetramer–tetramer interactions.<sup>30</sup> One of the major structural questions to be addressed is the precise role of the sequence extensions in the various  $\beta$ -crystallin assemblies, and from there to ascertain the structural consequences of their natural truncations by enzymes during lens development.

### Origins of crystallins

Knowledge of these structures has revealed much about the evolutionary origins of crystallins, which in turn can illuminate the present-day functional role of crystallins in the eye lens. The lens  $\beta$  $\gamma$ -crystallins have evolved or been recruited from ancient proteins, with their relatives turning up in unexpected places.<sup>31</sup> A two-domain monomeric version is found in a prokaryote and a one-domain molecule in a slime mould.<sup>32</sup> In the human genome a close relative of  $\beta$ -crystallins was found called Aim1, for absent in melanoma; it comprises 12 Greek key motifs and can be modelled as a domain-swapped trimer.<sup>33</sup> As more complete genomes are sequenced it may be possible to trace more relatives and establish context-specific functions for crystallin domains, perhaps confirming the idea that modern members of the  $\beta$  $\gamma$ -crystallins have evolved in the cellular response to changes in morphology.<sup>33</sup> Non-lens expression of

$\gamma$ S-crystallin, in cornea and retina, increases with age in young mice, suggesting it may be a general stress protein for the eye.<sup>34</sup>

### $\alpha$ -Crystallin and small heat-shock proteins

$\alpha$ -Crystallin is thought to function in the lens as a molecular chaperone by binding and stabilising early unfolding intermediates of other crystallins.<sup>35</sup> Recently the three-dimensional structure of a distantly related small heat-shock protein from the kingdom of the archaea has been solved.<sup>7</sup> The crystal structure shows that the C-terminal ' $\alpha$ -crystallin domain' folds into two  $\beta$ -sheets, so that superficially it resembles the  $\beta$  $\gamma$ -domain fold. However, the topology of the way the strands are organised is different and so it is unlikely they share a common ancestral origin. Pairs of the small heat-shock domains are woven together as one of the  $\beta$ -sheet strands is intercalated into the  $\beta$ -sheet of the partner domain, suggesting that the basic assembly unit is a dimer. These dimers further associate into a tetramer of dimers, which are further organised into trimers, with a total of 24 domains arranged into a sphere with cubo-octahedral symmetry. The sphere is pierced by several small windows and encapsulates a central cavity 65 Å in diameter. The N-terminal part of the polypeptide chain is invisible and presumed disordered inside the central cavity. This tightly bound oligomer functions as a molecular chaperone.

Previous multi-sequence alignments of all known members of the  $\alpha$ -crystallin/small heat-shock family had predicted quite accurately the regions of the sequence which formed  $\beta$ -strands,<sup>6</sup> and the actual topology of the  $\beta$ -sheets had been correctly predicted to be immunoglobulin-like.<sup>36</sup> It is very likely that the C-terminal domain of  $\alpha$ -crystallin will fold in a similar way. It is less certain how the higher assembly of  $\alpha$ -crystallin will compare with this archae structure. The average mass of  $\alpha$ -crystallin oligomers requires more than 24 subunits, and also requires a mechanism for generating a wide range of multimers of different sizes. What is fascinating about this small heat-shock protein structure is the observation that much of the disordered N-terminal domain is on the inside of the sphere, while in a plant small heat-shock protein, denatured substrate coats the outside.<sup>37</sup>

### Cataract: mutations

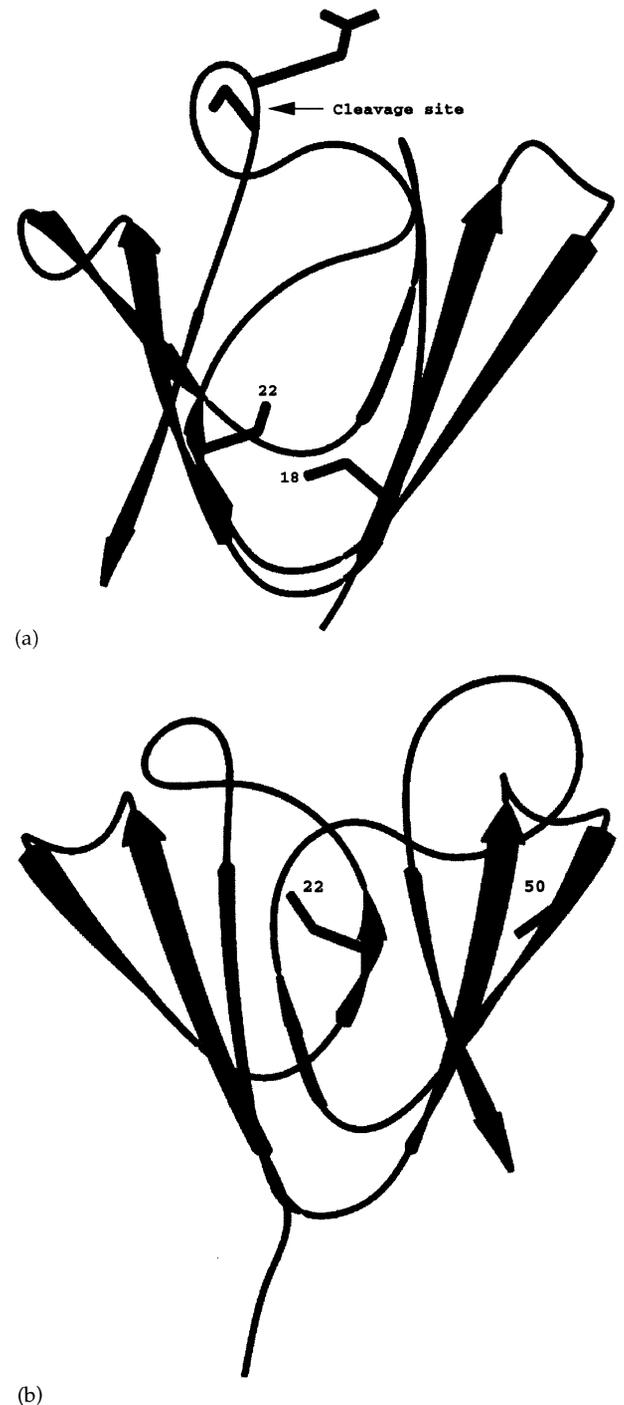
A few inherited cataracts in humans and mice have now been mapped to crystallin genes.<sup>38</sup> In order to interpret the effects of these mutations it is important to establish how much damaged protein is expressed and at what stage of lens development. Presumably if large amounts of unfoldable proteins are produced they will overwhelm the proteolysis machinery. The downstream effects of damaged protein on the viability of the lens cells at different stages of development are unclear. If the end result is no damaged protein in the cell, then the pure effect of a missing crystallin can be evaluated.

Hereditary human 'Coppock-like' cataract affects the central part of the lens formed during embryogenesis. A cluster of sequence changes was found in the vicinity of the promoter region of the pseudogene for  $\gamma$ E-crystallin which may result in its reactivation.<sup>39</sup> The pseudogene has an in-frame stop codon<sup>12</sup> and reactivation would result in high expression of an early protein truncated after Gln 54 in the middle of the **b** strand of motif 2 (Fig. 1). Such a protein will not fold and would lead to aggregation and precipitation. A mutation in human crystallin  $\beta$ B2, a gene that is highly expressed in the adult lens, caused autosomal dominant cerulean cataract with peripheral concentric opacities and occasional central radial lesions.<sup>40</sup> An affected homozygote had a single base change that created a stop codon in place of Gln 135, causing the polypeptide to be shortened by 51 residues (Fig. 1). If such a truncated molecule is highly expressed, it is possible that the N-terminal domain might fold, but the C-terminal domain will not. Removal of such a large chunk of the polypeptide would prevent the first motif of the C-terminal domain from folding, leaving a piece of peptide without the ability to bury its hydrophobic side chains in a regular way. It is expected that such a disabled molecule would self-aggregate. As normal  $\beta$ B2 is a subunit component of the majority of  $\beta$ -crystallin oligomers, suggested to have a solubilising role,<sup>41</sup> the knock-on effects of a disabled or absent molecule could be quite broad. Aggregating mutant  $\beta$ B2, or destabilised  $\beta$ -crystallin assemblies, may disrupt other  $\beta$ -crystallin molecules, perhaps over-burdening the chaperone capacity of  $\alpha$ -crystallin. An autosomal dominant congenital cataract has also been mapped to the gene for  $\alpha$ A-crystallin.<sup>42</sup> A missense mutation replaces arginine with cysteine in a conserved position equivalent to the seventh  $\beta$ -strand of the archaean  $\alpha$ -crystallin-domain structure.

An abnormal  $\beta$ B2 polypeptide chain has been found in the Philly mouse, a strain with an inherited cataract.<sup>43</sup> This results in the last  $\beta$ -strand of the C-terminal domain having an inappropriate sequence, including a bulky hydrophobic side chain in the position of the conserved serine of the fourth Greek key motif (Fig. 1).  $\beta$ B2 from the cataractous lens was found in the heavy-molecular-weight aggregates rather than the normal  $\beta$ -crystallin oligomers.<sup>44</sup> The mouse mutant Elo results in small eyes. The mutation appears to affect only the central fibre cells in early development and maps close to the early  $\gamma$ E-crystallin gene.<sup>45</sup> There is a single base pair deletion that would give rise to a truncated polypeptide in which the fourth Greek key motif was replaced by an eleven-residue extension of a different reading frame. As two Greek key motifs are required to form a domain, such a mutated protein would not be expected to fold properly and would be insoluble, sticking to other proteins in a random way.

A recent knock-out experiment linked connexins of gap junctions with a cytoplasmic crystallin.<sup>46</sup> Gap junction channels consist of two hemichannels called connexons located in the plasma membrane of two adjacent cells. Each connexon is composed of six

subunits called connexins. Two connexins are co-localised in lens fibre gap junction plaques ( $\alpha$ 3 (Cx46) and  $\alpha$ 8 (Cx50)). Disruption of the connexin 46 gene in mice led to a loss of protein solubility, proteolysis and nuclear cataract similar in morphology to human senile cataract. Nuclear  $\gamma$ -crystallin was cleaved in the N-terminal domain at a site consistent with the involvement of apoptotic proteases (Fig. 6a). A human autosomal



**Fig. 6.** Location of oxidising cysteines in the N-terminal domains of bovine  $\beta$ B2 and  $\gamma$ B-crystallins. (a) In  $\gamma$ B Cys 22 is close to Cys 18 and can readily form a disulphide. Note the exposed nature of the **cd** loop of motif 2. The arrow shows the  $\gamma$ -crystallin cleavage site associated with the cataract of the connexin 46 knock-out mouse. (b) In  $\beta$ B2 cysteine 22 is distal from cysteine 50, the only other cysteine in the  $\beta$ B2 polypeptide.

dominant 'zonular pulverulent' cataract has been mapped to the gene coding for connexin 50, resulting in the substitution of serine for a conserved proline.<sup>47</sup> The connexin story shows how defects in a membrane lens protein can cause quite unexpected damage to cytoplasmic crystallins and gives hints as to the role of apoptosis in damaged cells.

### Cataract: modifications

Cataract is mainly a problem of old age and so it is important to assess the effects of post-translational modifications on the long-lived crystallins. Global analyses of crystallin cleavage, oxidation and deamidation, as a function of age, are currently under way using electrospray mass spectrometry and highly resolving two-dimensional gel electrophoresis.<sup>27,28,48</sup> Specific cleavages, particularly in the  $\beta$ -crystallin N-terminal extensions, have already been shown to occur in the early stages of lens development, with the levels of truncation increasing with age. In the case of  $\beta$ B1, extensive truncations are associated with changes in oligomer size, and the truncated proteins remain in the water-soluble phase. Nearly all the water-soluble human crystallins from the adult lens have been subjected to deamidation resulting in more acidic proteins.<sup>28</sup> As there is a gradual increase in water-insoluble proteins with old age, similar analyses are required on these more difficult proteins.

Looking at the X-ray structures, truncation of the sequence extensions is the sort of change that might alter solubility/interaction properties of  $\beta$ -crystallins. Cleavages in the connecting peptide, leading to unpairing of domains, has already been shown to result in dramatic changes in solubility in the case of bovine  $\gamma$ B, in line with the observations of cleaved domains of human  $\gamma$ D found in cataract. The consequences of deamidation are likely to depend on location. Many changes would not be expected to destabilise the structure, but may affect solubility. Deamidation of extremely conserved glutamines, such as the topologically conserved Gln 54 and Gln 143 in motifs 2 and 4, which provide specific hydrogen bonds that help grip the two domains together (Fig. 3), may perturb domain pairing.

The  $\beta\gamma$ -crystallins are rich in sulphhydryl groups and these are distributed both in the cores and on the domain surfaces.<sup>5</sup> The surface cysteines may oxidise with other molecules and contribute to the formation of large complexes of diminished water solubility. Intramolecular oxidation could be relatively harmless, as in the observed disulphide bridge between Cys 18 and Cys 22 (Fig. 6a) in bovine  $\gamma$ B-crystallin.<sup>49</sup> The major human  $\gamma$ -crystallins,  $\gamma$ C and  $\gamma$ D, do not have such a pair although human  $\gamma$ S does.  $\beta$ B2 has a cysteine at a position topologically equivalent to Cys 22, but the only other one is at a position equivalent to Cys 50 (Figs. 1, 6b). Intramolecular disulphide formation of these cysteines would be expected to disrupt the structure. Our crystal structures show that these cysteines in  $\beta$ B2 have formed mixed

disulphides with low-molecular-weight reducing agent. These cysteines, however, form an intramolecular disulphide bond in nuclear cataract compared with an age-matched normal lens of a 60-year-old.<sup>50</sup> It is not known whether this molecular oxidation was the initiating event leading to a disrupted molecule or whether other events denatured the molecule first and subsequently led to oxidation.

### References

1. Wright AF. A searchlight through the fog. *Nature Genet* 1997;17:132-4.
2. Harding J. *Cataract: biochemistry, epidemiology and pharmacology*. London: Chapman & Hall, 1991.
3. Delaye M, Tardieu A. Short-range order of crystallin proteins accounts for eye lens transparency. *Nature* 1983;302:415-7.
4. Benedek GB. Cataract as a protein condensation disease. The Proctor lecture. *Invest Ophthalmol Vis Sci* 1997;38:1911-21.
5. Slingsby C, Norledge B, Simpson A, Bateman OA, Wright G, Driessen HPC, *et al*. X-ray diffraction and structure of crystallins. *Prog Ret Eye Res* 1997;16:3-29.
6. Caspers G-J, Leunissen JAM, de Jong WW. The expanding small heat shock protein family, and structure predictions of the conserved ' $\alpha$ -crystallin domain'. *J Mol Evol* 1995;40:238-48.
7. Kim KK, Kim R, Kim S-H. Crystal structure of a small heat-shock protein. *Nature* 1998;394:595-9.
8. Lubsen NH, Aarts HJM, Schoenmakers JGG. The evolution of lenticular proteins: the  $\beta$ - and  $\gamma$ -crystallin super gene family. *Prog Biophys Mol Biol* 1998;51:47-56.
9. Mayr E-M, Jaenicke R, Glockshuber R. The domains in  $\gamma$ B-crystallin: identical fold - different stabilities. *J Mol Biol* 1997;269:260-9.
10. Norledge BV, Mayr E-M, Glockshuber R, Bateman OA, Slingsby C, Jaenicke R, *et al*. The X-ray structures of two mutant crystallin domains shed light on the evolution of multi-domain proteins. *Nature Struct Biol* 1996;3:267-74.
11. Srivastava OP, Srivastava K, Silney C. Covalent modification at the C-terminal end of a 9 kDa  $\gamma$ D-crystallin fragment in human lenses. *Exp Eye Res* 1994;58:595-604.
12. Meakin SO, Du RP, Tsui L-C, Breitman ML.  $\gamma$ -crystallins of the human eye lens: expression analysis of five members of the gene family. *Mol Cell Biol* 1987;7:2671-9.
13. Broide ML, Berland CR, Pande J, Ogun OO, Benedek GB. Binary-liquid phase separation of lens protein solutions. *Proc Natl Acad Sci USA* 1991;88:5660-4.
14. Norledge BV, Hay RE, Bateman OA, Slingsby C, Driessen HPC. Towards a molecular understanding of phase separation in the lens: a comparison of the X-ray structures of two high  $T_c$   $\gamma$ -crystallins,  $\gamma$ E and  $\gamma$ F, with two low  $T_c$   $\gamma$ -crystallins,  $\gamma$ B and  $\gamma$ D. *Exp Eye Res* 1997;65:609-30.
15. de Jong WW, Lubsen NH, Kraft HJ. Molecular evolution of the eye lens. *Prog Ret Eye Res* 1994;13:391-442.
16. Palme S, Jaenicke R, Slingsby C. Unusual domain pairing in a mutant of bovine lens  $\gamma$ B-crystallin. *J Mol Biol* 1998;279:1053-9.
17. Basak AK, Kroone RC, Lubsen NH, Naylor CE, Jaenicke R, Slingsby C. The C-terminal domains of  $\gamma$ S-crystallin pair about a distorted twofold axis. *Protein Eng* 1998;11:337-44.
18. Palme S, Jaenicke R, Slingsby C. X-ray structures of three interface mutants of  $\gamma$ B-crystallin from bovine eye lens. *Protein Sci* 1998;7:611-8.
19. Bax B, Lapatto R, Nalini V, Driessen H, Lindley PF, Mahadevan D, *et al*. X-ray analysis of  $\beta$ B2-crystallin and evolution of oligomeric lens proteins. *Nature* 1990;347:776-80.

20. Mayr E-M, Jaenicke R, Glockshuber R. Domain interactions and connecting peptides in lens crystallins. *J Mol Biol* 1994;235:84–8.
21. Trinkl S, Glockshuber R, Jaenicke R. Dimerization of  $\beta$ B2-crystallin: the role of the linker peptide and the N- and C-terminal extensions. *Protein Sci* 1994;3:1392–400.
22. Hope JN, Chen H-C, Hejtmancik JF. Aggregation of  $\beta$ A3-crystallin is independent of the specific sequence of the domain connecting peptide. *J Biol Chem* 1994;269:21141–5.
23. Norledge BV, Trinkl S, Jaenicke R, Slingsby C. The X-ray structure of a mutant eye lens  $\beta$ B2-crystallin with truncated sequence extensions. *Protein Sci* 1997;6:1612–20.
24. Wieligmann K, Norledge B, Jaenicke R, Mayr E-M. Eye lens  $\beta$ B2-crystallin: circular permutation does not influence the oligomerization state but enhances the conformational stability. *J Mol Biol* 1998;280:721–9.
25. Wright G, Basak AK, Wieligmann K, Mayr E-M, Slingsby C. Circular permutation of  $\beta$ B2 changes the hierarchy of domain assembly. *Protein Sci* 1998;7:1280–5.
26. Slingsby C, Bateman OA. Quaternary interactions in eye lens  $\beta$ -crystallins: basic and acidic subunits of  $\beta$ -crystallins favour heterologous association. *Biochemistry* 1990;29:6592–9.
27. Ma Z, Hanson SRA, Lampi KJ, David LL, Smith DL, Smith JB. Age-related changes in human lens crystallins identified by HPLC and mass spectrometry. *Exp Eye Res* 1998;67:21–30.
28. Lampi KJ, Ma Z, Hanson SRA, Azuma M, Shih M, Shearer TR, *et al.* Age-related changes in human lens crystallins identified by two-dimensional electrophoresis and mass spectrometry. *Exp Eye Res* 1998;67:31–43.
29. Ajaz MS, Ma Z, Smith DL, Smith JB. Size of human lens  $\beta$ -crystallin aggregates are distinguished by N-terminal truncation of  $\beta$ B1. *J Biol Chem* 1997;272:11250–5.
30. Nalini V, Bax B, Driessen H, Moss DS, Lindley PF, Slingsby C. Close packing of an oligomeric eye lens  $\beta$ -crystallin induces loss of symmetry and ordering of sequence extensions. *J Mol Biol* 1994;236:1250–8.
31. Wistow G. Molecular biology and evolution of crystallins: gene recruitment and multifunctional proteins in the eye lens. Austin/New York: RG Landes/Springer, 1995.
32. Clout NJ, Slingsby C, Wistow GJ. An eye on crystallins. *Nature Struct Biol* 1997;4:685.
33. Ray ME, Wistow G, Su YA, Meltzer PS, Trent JM. 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* 1997;94:3229–34.
34. Sinha D, Esumi N, Jaworski C, Kozak CA, Pierce E, Wistow G. Cloning and mapping the mouse Crygs gene and non-lens expression of  $\gamma$ S-crystallin. 1998. <http://www.molvis.org/molvis/v4/p8>
35. Horwitz J. Alpha-crystallin can function as a molecular chaperone. *Proc Natl Acad Sci USA* 1992;89:10449–53.
36. Mornon JP, Halaby D, Malfois M, Durand P, Callebaut I, Tardieu A. Alpha-crystallin C-terminal domain: on the track of an Ig fold. *Int J Biol Macromol* 1998;22:219–27.
37. Lee GJ, Roseman AM, Saibil HR, Vierling E. A small heat shock protein stably binds heat-denatured model substrates and can maintain substrate in a folding-competent state. *EMBO J* 1997;16:659–71.
38. Hejtmancik JF. The genetics of cataract: our vision becomes clearer. *Am J Genet* 1998;62:520–5.
39. Brakenhoff RH, Henskens HAM, van Rossum MWPC, Lubsen NH, Schoenmakers JGG. Activation of the gammaE-crystallin pseudogene in the human hereditary Coppock-like cataract. *Hum Mol Genet* 1994;3:279–83.
40. Litt M, Carrero-Valenzuela R, LaMorticella DM, Schultz DW, Mitchell TN, Kramer P, Maumenee IH. Autosomal dominant cerulean cataract is associated with a chain termination in the human  $\beta$ -crystallin gene CRYBB2. *Hum Mol Genet* 1997;6:665–8.
41. Bateman OA, Slingsby C. Structural studies on  $\beta$ H-crystallin from bovine eye lens. *Exp Eye Res* 1992;55:127–33.
42. Litt M, Kramer P, LaMorticella DM, Murphey W, Lovrien EW, Weleber RG. Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. *Hum Mol Genet* 1998;7:471–4.
43. Chambers C, Russell P. Deletion mutation in an eye lens  $\beta$ -crystallin. *J Biol Chem* 1991;266:6742–6.
44. Russell P, Chambers C. Interaction of an altered  $\beta$ -crystallin with other proteins in the Philly mouse lens. *Exp Eye Res* 1990;50:683–7.
45. Cartier M, Breitman ML, Tsui LC. A frameshift mutation in the  $\gamma$ E-crystallin gene of the Elo mouse. *Nature Genet* 1992;2:42–5.
46. Gong X, Li E, Klier G, Huang Q, Wu Y, Lei H, *et al.* Disruption of  $\alpha$ 3 connexin gene leads to proteolysis and cataractogenesis in mice. *Cell* 1997;91:833–43.
47. Shiels A, Mackay D, Ionides A, Berry V, Moore A, Bhattacharya S. A missense mutation in the human connexin 50 gene (GJA8) underlines autosomal dominant 'zonular pulverulent' cataract, on chromosome 1q. *Am J Hum Genet* 1998;62:526–32.
48. Garland DL, Duglas-Tabor Y, Jimenez-Asensio J, Datiles MB, Magno B. The nucleus of the human lens: demonstration of a highly characteristic protein pattern by 2D electrophoresis and introduction of new method of lens dissection. *Exp Eye Res* 1996;62:285–91.
49. Najmudin S, Nalini V, Driessen HPC, Slingsby C, Blundell TL, Moss DS, Lindley PF. Structure of the bovine eye lens protein  $\gamma$ B( $\gamma$ II)-crystallin at 1.47 Å. *Acta Cryst* 1993;D49:223–33.
50. Takemoto LJ. Disulphide bond formation of cysteine-37 and cysteine-66 of  $\beta$ B2 crystallin during cataractogenesis of the human lens. *Exp Eye Res* 1997;64:609–14.
51. Evans SV. SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J Mol Graphics* 1993;11:134–8.