Version of Record: https://www.sciencedirect.com/science/article/pii/S0022283622001334
Manuscript_59f53ecb5ad3f9bcd5de89f677fccc87

Acquired disorder and asymmetry in a domain-swapped model for γ-crystallin aggregation

Vatsala Sagar and Graeme Wistow
Section on Molecular Structure and Functional Genomics, National Eye Institute, National Institutes of Health, Bethesda, MD 20892 USA

Corresponding Author:
Graeme Wistow, Ph.D.,
Section on Molecular Structure and Functional Genomics,
National Eye Institute, Building 6 Room 106,
National Institutes of Health, Bethesda, MD, USA.
Tel: 301-402-3452
Email:graeme@helix.nih.gov

© 2022 published by Elsevier. This manuscript is made available under the Elsevier user license
https://www.elsevier.com/open-access/userlicense/1.0/

Summary

Misfolding and aggregation of proteins occur in many pathological states. Because of the inherent disorder involved, these processes are difficult to study. We attempted to capture aggregation intermediates of γS-crystallin, a highly stable, internally symmetrical monomeric protein, by crystallization under mildly acidic and oxidizing conditions. Here we describe novel oligomerization through strained domain-swapping and partial intermolecular disulfide formation. This forms an octamer built from asymmetric tetramers, each of which comprises an asymmetric pair of twisted, domain-swapped dimers. Each tetramer shows patterns of acquired disorder among subunits, ranging from local loss of secondary structure to regions of intrinsic disorder. The octamer ring is tied together by partial intermolecular disulfide bonds, which may contribute to strain and disorder in the octamer. Oligomerization in this structure is self-limited by the distorted octamer ring. In a more heterogeneous environment, the disordered regions could serve as seeds for cascading interactions with other proteins. Indeed, solubilized protein from crystals retain many features observed in the crystal and are prone to further oligomerization and precipitation. This structure illustrates modes of loss of organized structure and aggregation that are relevant for cataract and for other disorders involving deposition of formerly well-folded proteins.

Introduction

Aggregation and deposition of formerly soluble proteins are processes associated with a wide range of diseases from dementias to cataract [1]. In some cases, proteins form relatively well-understood fibrils, as in amyloidoses. In this process there is a secondary structure transition in which ordered structures are lost and refolded into a different conformation, usually short β-strands, which form fibrils. In other cases, destabilization of structural elements by mutation or modification can reduce energy barriers to unfolding and exploration of new folds [2; 3]. However, the lack of stable structures, heterogeneity and the transience of different intermediate forms makes structural analysis of proteins on the path to unfolding a challenge, similar to that presented by intrinsically disordered proteins [4; 5].

Crystallins are long-lived, soluble proteins of the eye lens and are essential for the optical properties of the lens and for the maintenance of transparency over a lifetime [6; 7; 8]. However, particularly with age, crystallins can become insolubilized, forming light-scattering aggregates which lead to cataract. There are three major classes of crystallins in most vertebrates: α-crystallins, which are small heat shock proteins, and the evolutionarily related β- and γ-crystallins [7; 9]. γ-crystallins are highly symmetrical, two-domain monomeric globular proteins expressed at high levels in mammalian lens [7]. They seem to have evolved to maintain close, non-bonding interactions under extreme molecular crowding [10; 11; 12; 13] and recent results suggest that at physiologically relevant concentrations, non-bonded dimerization of γS-crystallins, which are monomeric under dilute conditions, can occur [14].

Several genetic cataracts in both humans and animal models are associated with γ-crystallin sequence variants [15; 16]. While crystallin aggregation in cataract generally appears to be amorphous, rather than ordered, in some model cases there is evidence for fibril formation [2; 17], while other forms of assembly, including crystallization have also been observed [15; 16]. Posttranslational modifications, such as oxidation, deamidation and racemization, in genetically normal γ-crystallins are also associated with cataract [6; 18; 19]. However, how these modifications lead to oligomerization and aggregation is not known. It has long been thought that intermolecular disulfide formation contributes [6; 20; 21; 22], while other models, such as partial domain swapping of secondary structure elements in a mimic of an oxidized γ-crystallin under low pH, have been proposed [23; 24; 25].

γS is the most abundant member of the γ-crystallin family in adult human lens [26] and is widely expressed in vertebrates [13; 27; 28]. Recent papers have described a crystallographic structure of a disulfide linked dimer of human γS monomers [29] and a molecular dynamics and NMR analysis of a cataract associated mutant of human γS that appears to have undergone domain-swapping [30].We explored the possibility of capturing unfolding/aggregation precursors of monomeric γS-crystallins under mild stress

by crystallization. Here we describe the results of one of these experiments which produced a crystal structure for a strained and asymmetrically disordered multimer that illustrates structural transformations that can occur for these proteins and suggests ways in which further aggregation could occur.

Results

Crystallization

In an attempt to capture possible precursors for cataract formation, we set up crystallization screens for γS-crystallins that were subjected to stresses including oxidation, mild acidity and mild denaturation by heating. Our previous analysis of chicken γS [13] showed that crystallization was strongly influenced by one residue. In chicken, L16 was found to be associated with avid crystallization, compared with Q16, a residue which is highly conserved in mammalian γ-crystallins. We tested L and Q16 variants for both chicken and mouse γS. Initial screens produced crystals for the L16 variant of mouse γS under mildly oxidizing and acidic conditions. Both native and selenomethionine derivative [31] crystals with the same space group grew from screening buffers at pH 5.0 with 0 or 0.1mM DTT from 1 week. Here we describe a selenomethionine derivative that grew without DTT and was analyzed at 3 weeks.

Structure determination

The crystal diffracted to ~2.9Å (Table 1). High B factors prevented direct phase determination and the structure was solved by molecular replacement [32], using the coordinates of separate N- and C-terminal domains taken from the NMR structure of mouse γS (PDB: 1ZWO). Although density was weak in several places, particularly for side chains, main chain connectivity was generally clear. The determined structure was very different from the familiar γ-crystallin monomer (Fig 1A), in which two similar domains (N and C) interact and are connected by a bent linker region. Instead, the crystal revealed an octameric ring of subunits (Fig 1B, C), with a core of relatively well-ordered N-domains arranged as two layers with four 4 N-domains in each layer. This core was surrounded by a halo of relatively disordered C-domains, with higher B factors on one side of the octamer than the other (Fig 1D). Interestingly, although C-domains of γ-crystallins generally appear to be more stable than N-domains, low pH preferentially destabilizes C-domain of γB-crystallin[33; 34].

Although the octamer formed from what were originally identical subunits, the asymmetric unit consists of four structurally non-identical subunits, with two asymmetric tetramers forming the octameric ring (Fig 1E, F). In the octamer, N-domains are stacked with two pairs of identical, symmetry related domains on opposite sides of the central ring, flanked by two pairs of non-identical subunits (fig 1E). For convenience we numbered the four non-identical subunits of the tetramer s1-s4, such that the octamer core consists of s1/s1; s2/s3; s3/s2 and s4/s4 stacked N-domains.

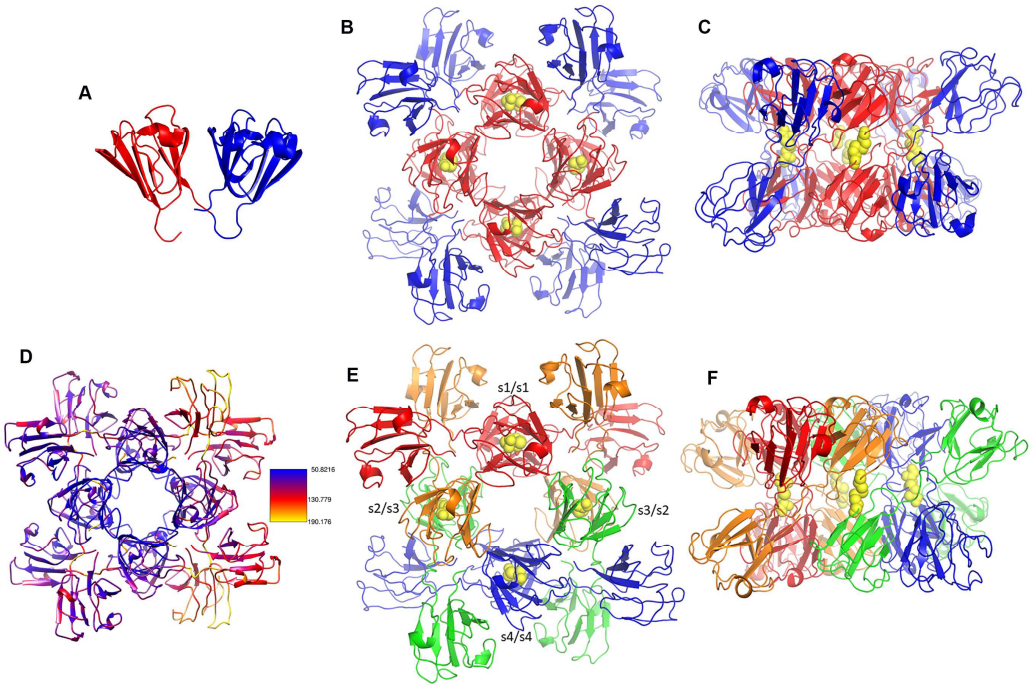


Figure 1. An asymmetric octameric structure for oxidized mouse L16-γS.

A) Ribbon structure of a typical monomeric γS-crystallin (PDB: 5VH1) showing N- and C-domains.
B) Arrangement of N (red)and C-domains (blue) and partial disulfide bonds (yellow) in the octamer (viewed down the axis of the N-domain assembly).
C) Orthogonal view of octamer showing two layers of domains and partial disulfides.
D) Heat map of B-factors for the octamer, showing asymmetry. The core is better folded than the periphery and one side is more disordered than the other. Color panel shows gradation of B factor values.
E) A view of the octamer colored by subunits (s1: red; s2: orange; s3: green; s4: blue) down the axis of the stacked N-domains.
F) Orthogonal view of the octamer colored by subunit as in E.

Domain swapping and disorder

In each subunit, density for the N-C domain linker or connecting peptide showed continuity consistent with an extended conformation (Fig 2A), rather than the bend seen in the classic monomer structure (Fig 1A). This results in domain-swapping [35], such that the N-domain of one subunit forms the conserved N/C domain interaction, characteristic of this superfamily [7], with the C-domain of another subunit (Fig 2B), forming s1/s2 and s3/s4 domain swapped pairs in the tetramer. This is the first crystallographic example of domain swapping in γ-crystallins. Crystals of the evolutionarily related, dimeric βB2-crystallin contain domain-swapped dimers (although this and other β-crystallins also form

different dimers without the domain swap [7; 36; 37; 38]). However, rather than the relaxed domain swap seen

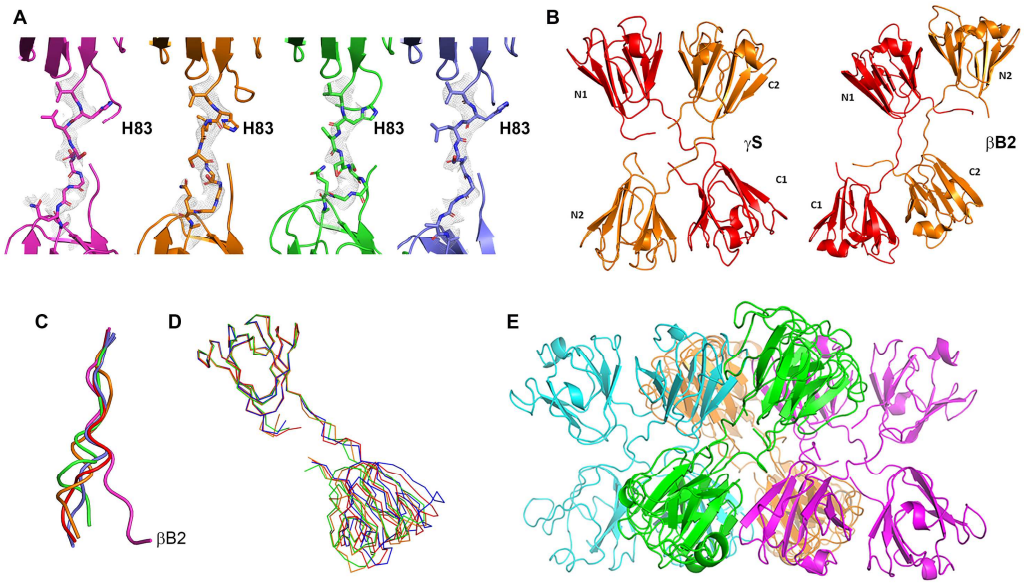


Figure 2. Domain-swapping in the octamer

A) Electron density ($2F_O$-F_C, 1σ) for the connecting peptides of s1-4. Side chains are poorly defined, but the direction and connectivity of the backbone is clear.
B) The domain arrangement for one the s1/s2 domain-swapped dimer of the γS octamer is compared with the relaxed domain-swap seen in crystals of βB2-crystallin. In the γS octamer, the dimer is twisted compared with the relaxed βB2 version.
C) Alignment of the main chain ribbon for the four connecting peptides (colored as in A) showing different degrees of twist and a shorter span than that in βB2 (violet).
D) Main chain traces for s1-s4 with N-domains aligned. The four C-domains have slightly different relative orientations.
E) Four domain swapped dimers (each pair colored differently) assemble to form the octamer.

in βB2, with roughly parallel connecting peptides (Fig 2B), the dimers in the γS octamer are twisted so that the linkers cross (Fig 2B). The linkers are also relatively shortened, compared with βB2, with a slight twist in the middle. In the relaxed human βB2 dimer the length of the connecting peptide (Cα-Cα for eight residues) (Fig 2B, C) is 20Å. In contrast, the four equivalent polypeptides in the γS octamer are more twisted, reflected by equivalent lengths of only16-18Å (Fig 2C).

Furthermore, each of the four linkers is slightly different in twist, with more similarity between the s1/s4 and s2/s3 pairs (Fig 2C). The conformation of some residues is ambiguous, suggesting the existence of multiple conformations. This linker differences are also reflected in differences in orientation of each C-domain relative to its N-domain (Fig 2D). Altogether, the octamer contains four domain-

swapped dimers (Fig 2E).

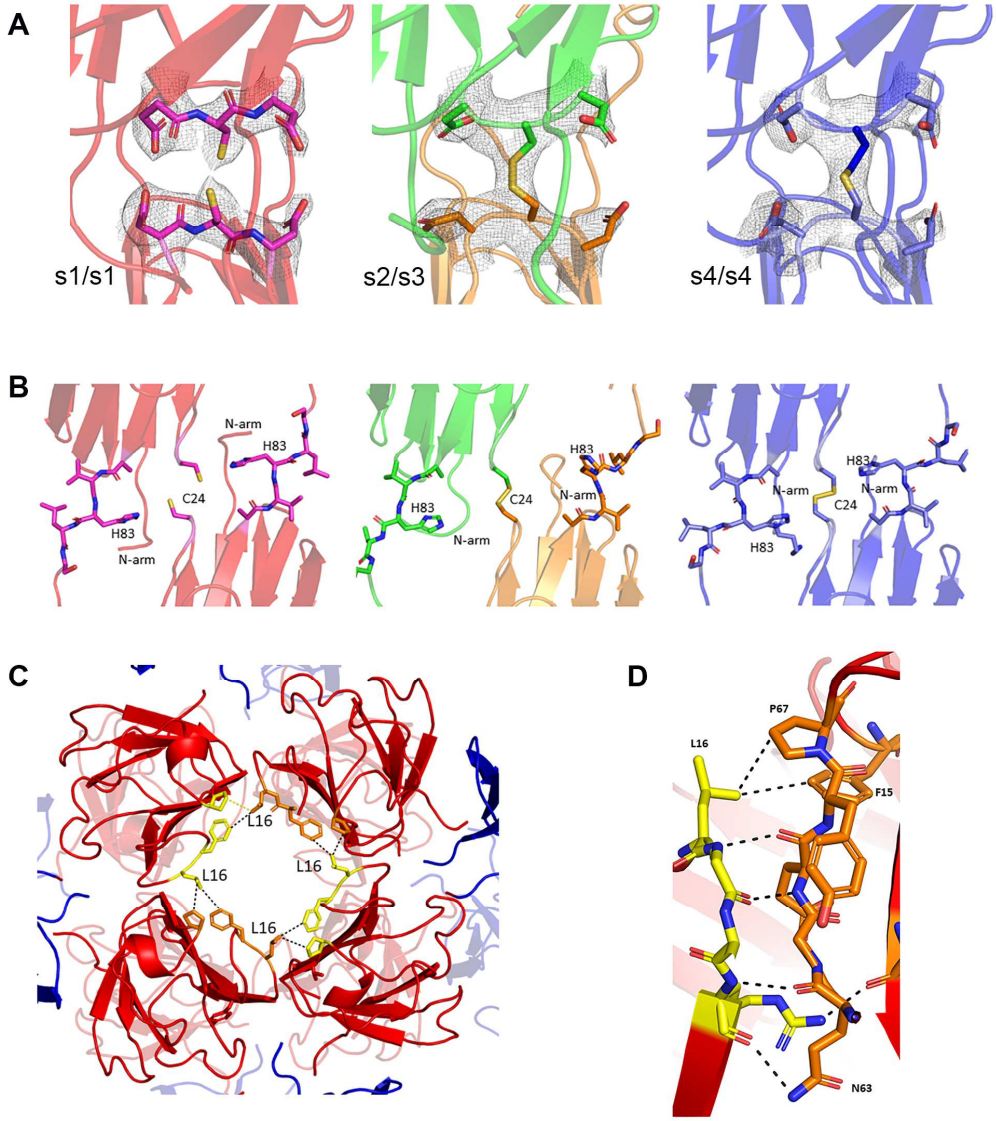


Figure 3. N-domain contacts in the octamer.

A) Electron density ($2F_O$-F_C, 1σ) for partial C24-C24 disulfide bond. The paired cysteines for s1 are close but do not form a disulfide.
B) The arrangement of the octamer layer interface. The C24-C24 pairs are flanked by H83 of the connecting peptides and the N-terminal (N-) arms.
C) Other N domain interactions. In each layer of the octamer, four N-domains form contacts centered on L16.
D) Details of interaction in C.

In addition to the domain swap, the layers of the octamer are held together by partial intermolecular disulfide bonds between residues C24 in pairs of N-domains which belong to different domain-swapped dimers. The electron density for the C24-C24 disulfide linking the N-domains of

subunits s2 and s3 and the symmetry related pair s4/s4 was unambiguous (Fig 3A). However, for the s1/s1 interaction, although the C24 side chains were very close, omit maps did not confirm a disulfide bond. There are two classes of disulfide in the octamer (s2/s3 and s4/s4) (Table 2). When compared with disulfides in PBD using the Disulfide Bond Dihedral Energy Server at UCLA (https://services.mbi.ucla.edu/disulfide/) the s2/s3 and s4/s4 disulfides were outliers in different ways. The s4/s4 bond was highly strained, with a ΔG of almost 21 kJ/mol and a highly unusual side chain torsional chi3 angle of -118.33. The s2/s3 bond also had unusual angles which do not fit into previously described categories [39]. The apparent failure to form the s1/s1 disulfide may be the result of strain across the ring caused by the strained bonding of s4/s4. Indeed, the formation of disulfides, pulling the two layers of the octamer together may be the cause of the twist in the domain linkers. This could propagate through the first strand of motif III of the C-domain and contribute to the disorder and asymmetry observed across the octamer structure (Fig 1D).

The C24-C24 contacts, whether disulfides or not, form the cores of four bundles of structural elements (Fig 3B). This bundle includes the N-C domain linkers, with the side chains of H83 (the residue which marks the start of the linker) close to, but not contacting, C24. The involvement of H83 in this structural unit may be significant, since crystallization occurred under a pH allowing histidine ionization. Mutagenesis of this and other residues may shed further light on the mechanism of domain swap. The short N-terminal arms (N-arm) of each subunit also form part of this bundle, lying roughly parallel to the C24-C24 axis, close to the N-domain of the paired subunit. This is reminiscent of the crystal structure of monomeric chicken γS [13] in which the N-arm also forms a (different) intermolecular contact suggesting that, in the lens, the N-arm may be able to participate in multiple intermolecular interactions.

In both layers of the octamer each of the four N-domains has contacts centered on L16, mainly involving main chain hydrogen bonds (Fig 3C, D). As estimated using PISA [40], the solvation energies (ΔG_{solv}) of these N-domain interactions range from only -0.2 to 2.2 kCal/mol, suggesting no major contribution to dimerization.

Interactions and disorder

Compared with a monomeric γ-crystallin, the octamer is highly disordered and the degree of disorder throughout the octamer correlates with the environment of each domain. The core N-domains have interfaces with each other and with domain swapped C-domains (Fig 1) and their main chain B factors range from 50-125Å^2 (compared with 10-40Å^2 for both domains in the monomeric chicken γS structure) while in the exposed C-domains B factors range from 60-200Å^2 (Fig 1D). Disorder is also reflected in the RMSD for each domain compared with the equivalent (N- or C-) for the crystals structure of chicken γS-crystallin (Table 3) with larger deviations in the C-domains, particularly for s4.

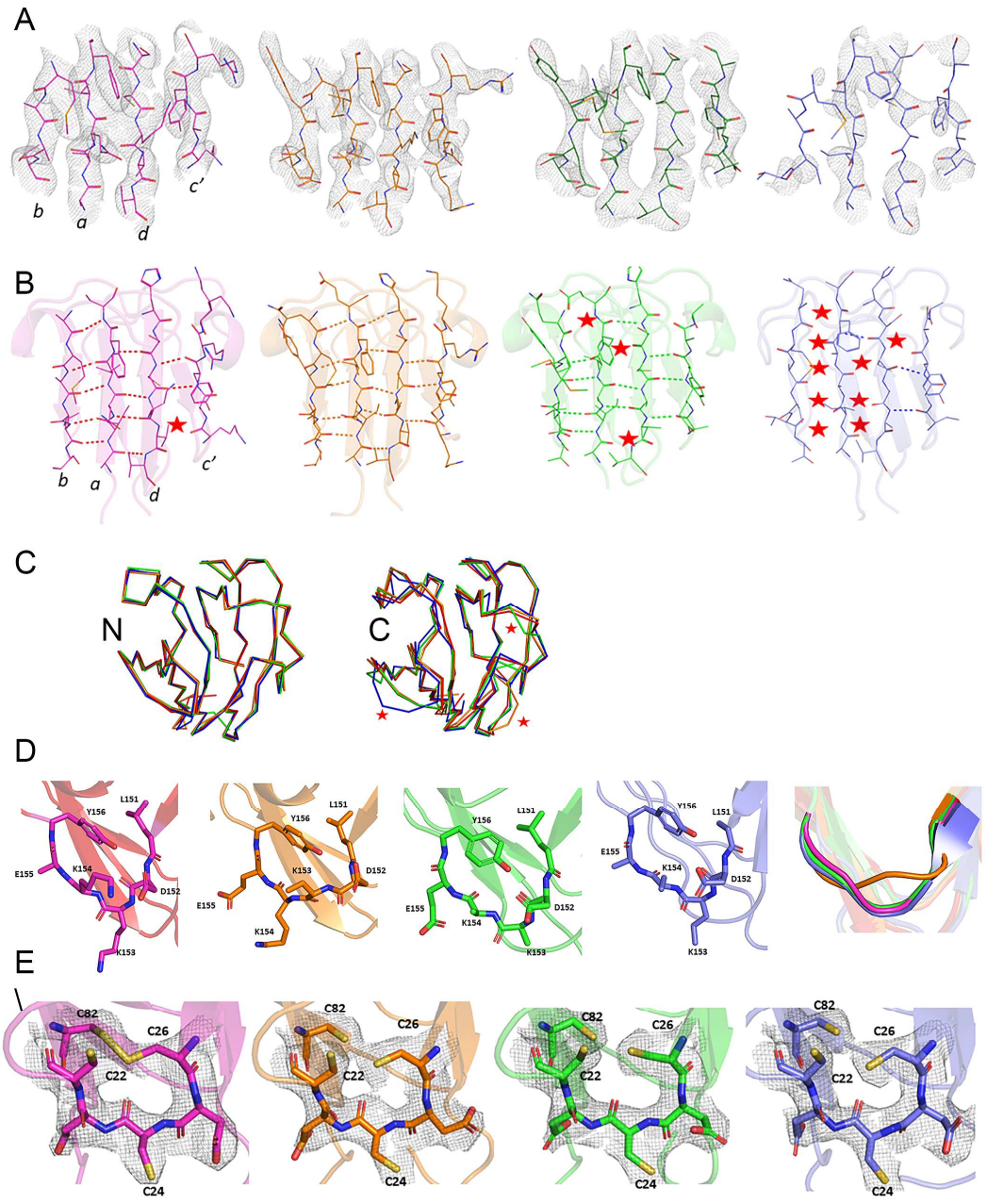


Figure 4. Secondary structure changes and disorder in the octamer.

A) Electron density for the surface exposed β-sheet of the C-domain in s1-s4. The *b-a-d-c'* order of β-strands is shown for s1. At 1σ density, several side chains are missing as well as parts of backbone in s4.

B) Apparent loss of hydrogen bonds in the exposed C-domain β-sheets in s1-s4. Asterisks indicate lost hydrogen-bonds, as judged by distance and geometry or by loss of density. Viewed from a slightly different angle from A) to make sheet structure clearer.

C) Alignment of chain trace for N- and C-domains of s1-s4 (colored as before) showing differences in the backbone path. Some major loop shifts are indicated by stars.

D) Loss of a Tyr corner in the C-domain of s2. Left: The expected arrangement in s1; right: main chain deviation in s2. Asterisk shows position of K153 which is either up or down depending on whether the bend is V-shaped or broadened.
E) Electron density ($2F_O$-F_C; 1σ) for C22, 24, 26, 82 in s1-4 (colored as in other figures) showing evidence for partial C22-82 disulfide.

In monomeric γ-crystallins, each domain is formed by a pair of 4-stranded (*a-b-c-d*) Greek key motifs, for a total of four motifs per monomer (I-IV) [41]. These motifs form paired β-sheets in each domain with three strands (*b-a-d*) coming from one motif and the fourth (*c'*) from the paired motif. The intramolecular interface between N- and C-domains consists principally of motifs II and IV which are consequently mostly buried, while the sheets formed mainly by motifs I and III are more exposed. In the octamer structure, the exposed β-sheets in the C-domains (mainly motif III) show much more disorder than those in the N-domains, reflected in weaker main chain and side chain density, particularly in s4 (Fig 4A). This disorder is also associated with loss of geometry and increased lengths of hydrogen bonds compared with canonical β-sheet (Fig 4B).

Throughout the βγ-crystallin superfamily, strands *a* and *b* of each motif form a characteristic β-hairpin connected by a highly structured loop [7; 41], stabilized by interactions with the *c-d* loop and strand *d*. In the octamer, the loss of structure for strand *b* in the conserved hairpin, particularly in motif III of subunit s4, is associated with a significant shift in the conserved surface loop that connects strands *a* and *b*. Indeed, across subunits s1-s4 the positions of surface loops in the C-domains shift (Fig 4C). In solution, the disorder of surface loops seen in the octamer could make the structure vulnerable to further non-native intermolecular interactions [1].

One contribution to the higher disorder of motif III in the octamer could be strain from twisting of the interdomain connecting peptide, resulting from the domain swap and relative compression of the chain resulting from the disulfide bonds pulling the layers together (Fig 2, 3). Tension from the twisted connecting peptides could propagate into strand *a* of motif III.

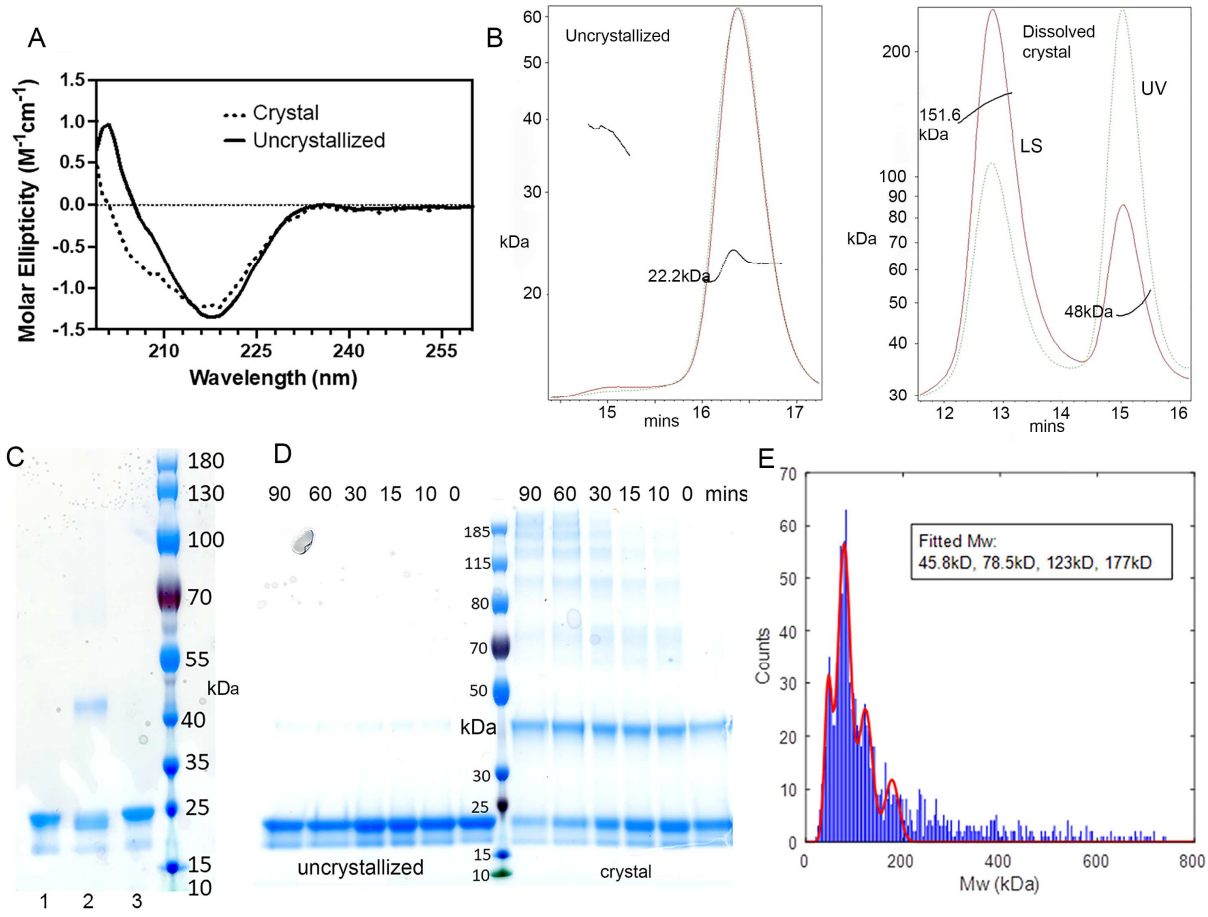
Another characteristic of βγ-crystallin domains is a tyrosine corner [42] in the second motif of each domain [6]. In the C-domain of subunit s2, this feature is lost, as illustrated by the shift in positions of K153 and K154 (Fig 4D). Although subunits s1/s2 are generally less disordered or distorted than s3/s4, the strain apparent in the octamer has different effects in each subunit and all are affected. Imperfect planarity in the octamer may cause specific strains that exert asymmetric effects in each octamer, affecting overall disorder and conformational differences among subunits.

Disulfides

In addition to the partial intermolecular C24-C24 disulfide, three other residues, C22, C26 and C82, are in positions where intramolecular disulfides are possible. Indeed, the equivalent of C22-C26 was

observed in aged crystals of bovine γB(II) [43] and also in the human γS-dimer [29]. In the octamer, there is density consistent with at least partial formation of C22-82 (Fig 4E), particularly in subunit s1. However main chain geometry is ambiguous in this bend region, suggesting that different conformations may occur. The C-domains also contain two cysteines that are partially (C114) or completely (C129) buried. Despite their normal inaccessibility, these residues have been found to be oxidized in human cataract [23]. Loss of tight folding and increased mobility of structures in the C-domain, as seen in the octamer, would allow increase exposure of these residues.

Figure 5: Multimerization and disorder in dissolved octamer crystals



A) CD spectra of uncrystallized monomer and dissolved crystals of γS-L16. Evidence for a loss of organized secondary structure in the crystal as the β-sheet minimum broadens with a shoulder at 205nm. Spectra were acquired at room temperature in 0.3 mg/ml in 50 mM sodium acetate pH 5.0, buffer using a 1-mm path length cuvette. Crystal samples were diluted tenfold (from 30% glycerol stock) and both spectra were acquired in 3% glycerol (v/v), which was added to the uncrystallized sample.

B) SEC-MALS of uncrystallized (left panel) and dissolved crystals of γS-L16 (right panel) at 1 week. Red trace shows light scattering (LS); green shows UV absorbance at 280nm. Monomer

shows a single major peak consistent with a size of 22kDa. Dissolved crystal shows two major peaks consistent with dimer (48kDa) and octamer (152kDa) size ranges in solution. Molecular mass scale is shown as kDa.

C) Non-reducing SDS PAGE of crystallized protein gives a 44kDa band consistent with a disulfide linked dimer. Lane 1 shows uncrystallized protein; 2 shows dissolved crystal under non-reducing conditions; 3 shows dissolved crystal under reducing conditions. The minor band apparently running below 20kDa in both native and crystallized protein was analyzed by mass spectroscopy and determined to be full-length protein containing an internal disulfide bond (as seen in figure 4E).

D) Non-reducing SDS PAGE of a time course (0-90 mins) of glutaraldehyde cross-linking of uncrystallized monomer and dissolved crystals of γS-L16. For uncrystallized protein (left side of gel) monomer is predominant. In contrast, the dissolved crystal shows patterns consistent with increasing multimer formation up to octamer.

E) Mass photometry (ISCAMS) of dissolved crystal solution. The MP data was plotted as kernel density estimate (KDE) distributions (blue histograms) with Gaussian function fit (red) for average molecular mass of each distribution component. Mass estimates are shown. This technique is performed at relatively low concentration which favors dissociation of the larger complexes.

Solution of crystals

So far, attempts to observe the octamerization process in solution have been unsuccessful: without the conditions for crystallization, aggregation leads to precipitation. The octamer structure may require the scaffold of the crystal to form size-limited multimers without forming precipitates. However, we were able to reverse the process and to solubilize octamer crystals. For 3-week-old crystals, addition of 30% glycerol gave more complete solubilization and the glycerol was diluted in subsequent steps. For consistency, the same protocol was used for crystals from 1-3 weeks. This process produced a protein solution which revealed similar properties to those seen in the crystal.

The far-UV CD spectrum of dissolved 3-week crystal (Fig. 5A) exhibited a shoulder/ minimum in molar ellipticity at ~ 205 nm which is indicative of some disordered structure relative to the CD spectrum of the monomer (which is very β-sheet in character, as expected, in having a broad ellipticity minimum at ~ 218 nm). Using the CAPITO online server, predicted secondary structure for the monomer (NRMSD of 0.71) was 51% β-strand, 18% α-helix and 46% irregular structure (random coil). In contrast, the dissolved crystal (NRMSD of 0.36) gave 44% β-strand, 10% α-helix, and 52% irregular structure. This is consistent with the loss of secondary structure seen particularly in the exposed regions of the C-domains of the octamer.

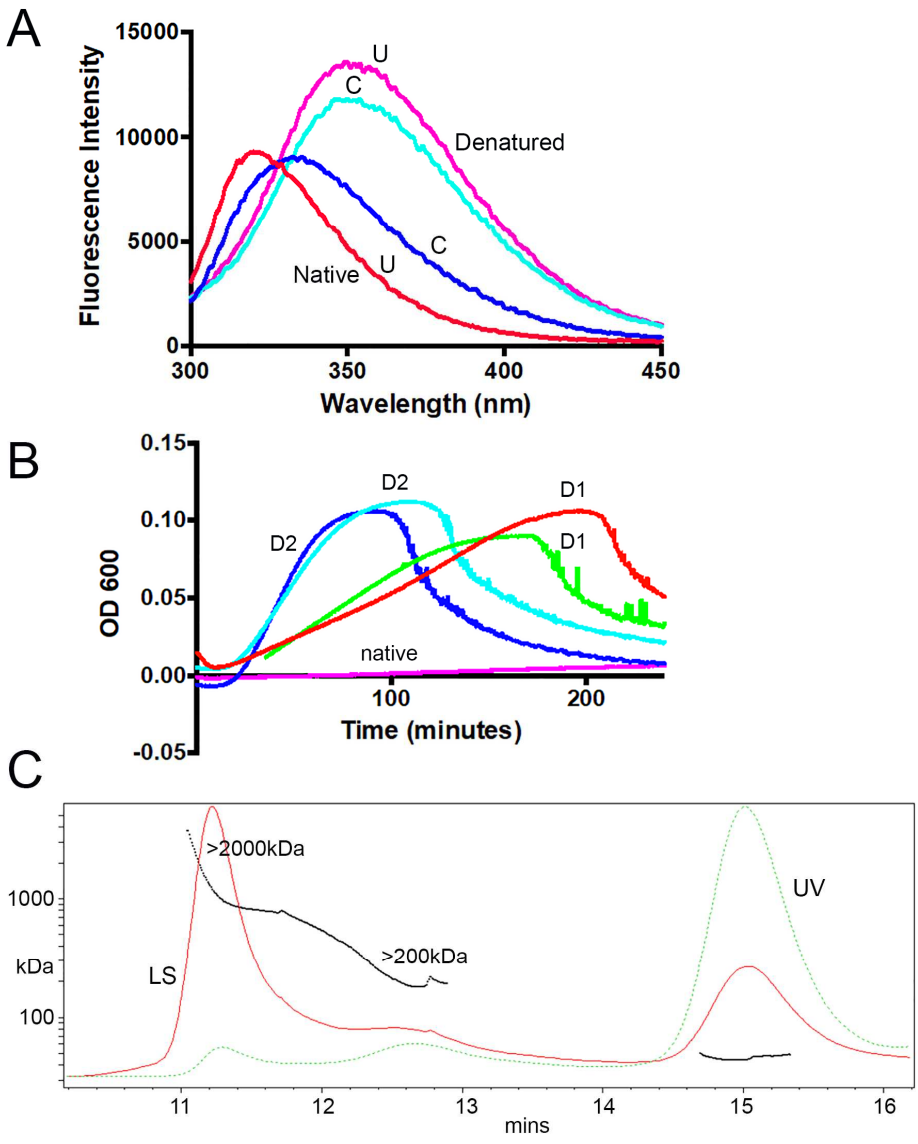
Dissolved crystals at 1 week were examined by SEC-MALS (Fig 5B) and gave a majority mass of 151,600 Da in solution, close to that (168 kDa) predicted for the octamer, while the uncrystallized protein was almost entirely monomer. The detection of likely octamer suggests that much of the crystal structure is retained in solution and also supports the existence of the domain-swap in forming the

octamer, since an octamer consisting of two layers of non-domain-swapped tetramers would have nothing to hold it together outside the crystal lattice. In the absence of cross-linking, the dissolved 1-week crystals resolved as monomers and dimers on SDS PAGE gels under non-reducing conditions, consistent with the partial intermolecular disulfide formation in the crystal (Fig 5C). After mild cross-linking with glutaraldehyde and resolution by SDS PAGE, uncrystallized protein gave only monomers while dissolved crystals gave bands corresponding to monomer, dimers and larger multimers (Fig 5D). A minor band appeared in SDS PAGE of both crystallized and uncrystallized protein. Mass spectroscopy of tryptic and chymotryptic digests confirmed that this minor component is full-length protein containing an internal disulfide between Cys82 and the peptide containing C22C24C26, consistent with the evidence of partial disulfide formation shown in Fig 4E. This bond is evidently relatively resistant to reduction in SDS PAGE buffer and its retention causes a mobility shift in the gel.

Samples of the dissolved crystals were also analyzed by mass photometry [44] (ISCAMS). This showed a mass of 177 kDa (consistent with octamer), 123 kDa (~hexamer), 78.5 kDa (~tetramer) and 45.8 kDa (~dimer) (Fig 5E).

Solutions of the dissolved 1-week crystals were protected with EZ-Link™ Maleimide-PEG2-Biotin in 8M urea and run on SDS PAGE. The dimer band (equivalent to Fig 5C) was excised, digested with trypsin and analyzed by liquid chromatography MS/MS. The MS raw files were analyzed by using the StavroX3.6.6 software package to identify disulfide bond cross-linked peptides. The cysteine cross-linked peptides through disulfide bonds were identified as summarized in Table 4. The C24-C24 disulfide-bond was dominant, while others (C22-C22, C22-C24, C22-C26, C24-C26, C26-C26) were possible but less likely. (This partial ambiguity is due to Asp-Cys repeats in the tryptic peptide "DC22DC24DC26"). The results confirm that dimerization was mediated through this peptide of the N-domain, most likely through C24-C24, as seen in the crystal.

Fig 6. The crystallized protein is aggregation-prone.



A) Trp fluorescence spectra of native protein and dissolved crystal (2 weeks old) in buffer and in 7M GdnCl. Unfolding the native protein red-shifts the maximum intensity. The dissolved crystal gives a red-shifted spectrum consistent with partial unfolding. Uncrystallized (U) and crystallized (C) samples in buffer are shown by red and blue lines respectively; the same preparation in 7M GdnCl are shown in pink and light blue.

B) Turbidity of solutions at 600nm. While uncrystallized monomer remains stable in solution at 60C, the dissolved crystal rapidly forms light scattering particles and precipitates under the same conditions. Results are shown for two preparations of crystal, both freshly dissolved (red and green traces, day (D) 1) and after a further 24 hrs at 20C, the same as the crystallization conditions (dark and light blue traces, day (D) 2), Turbidity increases more rapidly on day 2 (D2). All samples contained a total of 3% glycerol (v/v).

C) SEC-MALS of 3-week-old dissolved crystals. Compared with data in Fig 5B, there is evidence for an increase in formation of larger species, up to $4x10^6$ Da, consistent with increasing tendency towards aggregation. Light scattering (LS) and UV (280nm) traces are indicated. Calculated size ranges for LS peaks are shown in kDa.

Overall, the results suggest that dissolving the crystals yielded protein in solution with similar characteristics to that what was seen in the crystal. Loss of tight folding in the crystallized protein was also demonstrated using tryptophan fluorescence (Fig 6A). Under non-denaturing conditions, Trp fluorescence in native γS-crystallin is quenched. In the presence of denaturant (7M guanidinium chloride) the protein unfolds, and the fluorescence increases in intensity and its maximum is red shifted. The dissolved 3-week crystal shows a red-shift even under non-denaturing conditions, consistent with a more exposed core structure. In denaturant, the spectra for both crystallized and uncrystallized proteins are similar, consistent with complete unfolding.

The thermal stability of the dissolved crystals was then tested by comparing solution turbidity with uncrystallized monomer (Fig 6B). At 60C, the monomer was stable for at least 3 hours, consistent with previous observations. In contrast, freshly dissolved and diluted 2-week crystals rapidly became turbid under the same conditions, increasing to a maximum after 2 hours which then declined as aggregates precipitated out of solution. Duplicate preparations gave similar profiles when freshly dissolved (D1) and also after a further 24 hrs (D2), however the D2 solution gave a more rapid increase in turbidity, suggesting that the processes of aggregation had advanced.

SEC-MALS of dissolved 3-week crystals also showed an increase in the size and abundance of larger multimers (Fig 6C) compared with the 1–week preparation (Fig 5B), suggesting that even in the crystal, aging is associated with a higher propensity for aggregation when in solution. Overall, these results are consistent with the loss of organized secondary structure and increased disorder in the crystallized octamer leading to further interactions and aggregation in solution.

Discussion

Protein aggregation and deposition is a common feature of diseases, particularly those related to aging [1; 45]. Inherited sequence variants may confer a predisposition to formation of non-native assemblies, but even stable, well-folded proteins can lose stability and begin to explore other conformational states through post-translational modifications such as oxidation. In most cells, active turnover of proteins mitigates against the accumulation of folding/unfolding variants and, in dividing cell populations, cell death can remove those cells that become compromised and threaten tissue function [46]. However, in terminally differentiated cells that experience no or low turnover, such as lens fiber cells and many neural cell types, age-related accumulation of pathological protein deposits may occur [47]. Cataract in the lens is

an example of this sort of process in which accumulated protein deposits impair vision [6; 18; 48]. Oxidation has long been implicated as an important component in cataractogenesis [6; 20]. While fibrillar structures have been identified in some crystallin aggregates, mechanisms of non-fibril aggregations are poorly understood.

Using the mild stress condition of pH 5 under non-reducing conditions, we were able to use crystallization to capture the octameric structure, derived from monomeric mouse γS-L16, presented here. Although L16 itself is involved in a domain contact, the energetic component of the total contact is small (-0.2 to 2.2 kCal/mol) and L16 is only part of the contact. γS-crystallins in mammals are highly conserved and the human ortholog is likely to be capable of similar reorganization. This gives insights into some of the processes that may occur during aggregation and oxidation of crystallins. More generally, this structure provides a new view of the kind of conformational changes that can occur in well-folded, globular proteins. Some initial modification, such as domain-swapping or disulfide formation, may promote a cascade in which disorder increases and other proteins are recruited Indeed, a recently published study has described a non-domain swapped C24-C24 dimer of human γS-crystallin and has shown evidence that this too has an increased propensity for unfolding and aggregation [29].

The octamer is an unusual example of a crystallized multimer, derived from identical monomers, which has acquired asymmetry. The asymmetry appears to be the result of strain caused by the geometric constraints of domain swapping and disulfide formation. The increased disorder is reflected in distortion of secondary structures, such as β-sheets and hairpin loops, that are well-defined in canonical, monomeric γ-crystallins. Distortions vary in different subunits in a consistent way across the tetrameric asymmetric unit, perhaps a result of the order in which the octamer is built up from dimer to tetramer to octamer. It may be that closing the "circle" of the octamer causes more stress at one side of the ring than the other. One result of this seems to be the failure to form one of the possible intermolecular disulfide bonds (s1/s1) on the opposite side of the central ring from the highly strained s4/s4 disulfide. Disorder varies considerably from one subunit to another, leading to regions with high levels of acquired intrinsic disorder, particularly in the C-domain of subunit s4.

For intrinsically disordered proteins, interactions with ligands may stabilize and organize secondary and tertiary structure elements [4]. Something like this is apparent in the octamer in which the regions of lowest B factors have the closest intermolecular contacts, as in the core of N-domains, while the most disordered regions, the C-domains, are more exposed. However, whether closer packing stabilizes otherwise disordered structures, or whether the better folded regions are more permissive for closer packing is not clear.

In the crystal, oligomerization has been self-limited by octamer formation. In solution and in the heterogeneous environment of the lens many other interactions could occur in a cascade leading to increased aggregation. The loss of organized structure apparent in the C-domains in the octamer could allow conformational changes that would expose buried cysteine residues, increasing the chances of further disulfide formation [22; 23]. In solution, the surface loops that show increased disorder in the octamer would be available to form intermolecular interactions of various kinds, perhaps including the kind of stacking of β-strands that occurs in fibrils [45].

Indeed, the behavior of protein from solubilized crystals is consistent with this idea. The crystals can be dissolved to produce a protein solution which has very similar characteristics of multimerization, loss of secondary structure and disulfide formation to those seen in the crystal. While the monomer is stable in solution even at 60C, the crystal solution rapidly becomes turbid and precipitates. Older crystals have a higher propensity for aggregation.

This structure illustrates ways in which well folded globular proteins can progress to aggregates as sequential events of linkage or conformational reorganization occur, each lowering the energy barrier to another step in a cascade. The high symmetry of γ-crystallins may have allowed us to capture this structural variant, but our results offer the possibility that other proteins on track for aggregation could be captured in a similar way.

Materials and Methods.

Protein Expression

The L16 version of full-length mouse γS-crystallin was prepared by site-specific mutagenesis of the previously described native sequence expression clone in the pET17b vector (EMD Millipore, Darmstadt, Germany) using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies, Santa Clara CA) and was expressed as described before [28]. For selenomethionine derivatization we used M9 selenomethionine (SeMET) growth media kits (Medicilon, Shanghai, China). For both versions, pellets were collected by centrifugation at 5,000 rpm. for 30 min and stored frozen at −80° and protein was purified as described before [13].

Crystallization

Purified protein from stock (5mg/ml in 50 mM Tris pH 7.4) was crystallized by the sitting-drop vapor-diffusion method using commercially available crystal screens: Index (Hampton Research, Aliso Viejo, CA), PEG/Ion and PEG/Ion 2 (Hampton Research), and Pi-minimal Screen (Jena Bioscience, Ann Arbor MI). A total of 288 conditions were tested using 1μL protein and 1μL screening buffer per drop.

Screens were incubated at 20°C. The condition producing the largest, most uniform crystal (0.15 M Sodium acetate pH 5.0, 22% PEG MME 550, 8% PEG 300) for L16 was selected for optimization of PEG concentration. Final conditions used 15% PEG MME 550 and 8% PEG 300). Crystals appeared at 1 week and were harvested at 1, 2 or 3 weeks.

X-Ray Analysis

3-week crystals were flash-frozen in mother liquor containing 20% ethylene glycol. Diffraction data were collected remotely at the NE-CAT 24ID-E beamline at the Advanced Photon Source (APS) at Argonne National Laboratory, Lemont, IL. Diffraction data were collected at 100K with an Eiger 16M detector. X-ray data collection and refinement statistics are shown in Table 2. Data were collected out to 2.9 Å. A native data set of 360 frames was collected at Se-edge of 12.66 keV with 10% transmission, 1° rotation and 1s exposure. The data were integrated and merged using HKL2000. R_{free} was monitored by using 5% of the reflections as a test set. The structure was solved by molecular replacement using the NMR structure of mouse γS crystallin (PDB 1ZWO) searching with N terminus and C terminus domains separately. Molecular replacement was performed using the Phaser MR program [32] from the Phenix crystallography software suite36 [49]. The Z-scores were RFZ=2.8 LLG=6394 TFZ==4.2. Structure was refined with TLS with four monomers using Phenix.Refine [49]. Variant regions, including disulfides, connecting peptides and surface loops were verified by omit maps, deleting resides and running three cycles of refinement. For the most disordered domain (C-domain of s4) molecular replacement with the equivalent domain of s2 gave a poor fit initially. The fit was improved by manual building and multiple rounds of refinement. To validate this structure, the coordinates of s2 were placed into the final density of s4, followed by 3 rounds of restrained refinement. R_{work}/Rf_{ree} rose to 0.3086/0.2163, and s2 atoms moved into s4 density

The structure is deposited as **PDB ID 7RJ0**.

Informatics

Figures were created using PyMOL (Schrödinger LLC, Cambridge MA). Structure alignments to calculate root-mean-square deviation (RMSD) were performed using Align in PyMOL. The DALI server was used to search for the closest structural alignment to s1 of the octamer in PDB. Contact energy was calculated using PISA [40].

Dissolving crystals

Octamer crystals were placed into 24 µL of 50 mM sodium acetate, pH 5.0 with 24 µL of 60% glycerol (30% v/v) and incubated overnight at room temperature. While fresh crystals did not need glycerol, for the age required to match the crystal structure, glycerol gave more complete solution and was used for different ages for consistency. Samples were centrifuged at 13,200rpm for 30 minutes and

the supernatant transferred to a fresh microcentrifuge tube. The concentration of protein was determined by nanodrop and diluted with 50 mM sodium acetate pH 5.0 to the appropriate concentration. Subsequent steps diluted the glycerol concentration tenfold or lower for different techniques. At each stage, the uncrystallized protein control was from the same original preparation used for crystallization and held at the same temperature, 20C.

Circular Dichroism

Circular dichroism (CD) spectra were recorded at room temperature on a Q100 spectropolarimeter (Applied Photophysics., Leatherhead, Surrey, UK) with spectral range from 163-1150nm. Far UV CD experiments were performed at protein concentration of 0.3 mg/mL in 50 mM sodium acetate pH 5.0, 3% glycerol buffer using a 1-mm path length cuvette. Data were collected over the wavelength range of 260–200 nm at 1-nm intervals. Ten scans were averaged and normalized by subtracting the base line recorded for the buffer. The results were plotted as molar ellipticity (θ) versus wavelength (nm).

SEC-MALS

Size Exclusion chromatography for multiple angle laser light scattering (SEC-MALS) measurements were performed at 25 °C using a Dawn HELEOS-II MALS with QELS DSL from Wyatt Technology (Santa Barbara, CA), with Agilent (HP) 1100 series pumps, Agilent autosampler, diode-array UV/Vis and the inline fluorescence detector. Running buffer was 50 mM sodium acetate pH 5.0, 50 mM NaCl. Flow rate was 0.4 ml/min. Column fractionation range was 500 -150,00 g/mol. The refractive index detector was the Optilab T-rEX from Wyatt Technology. Data acquisition and evaluation was performed using Astra from Wyatt Technology. Bovine serum albumin (Sigma, St Louis, MO) was used to normalize detectors. Molar absorptivity (280nm) of 1.9 M-1cm-1 was used to calculate concentration. Separations were carried out using a SEC Analytical Column, 5µm 150Å 7.8mm ID, Standard (Wyatt Technologies, WTC-015S5.). The buffer used for gel filtration was 50 mM sodium acetate pH 5.0, 50 mM NaCl and the solution concentration was 0.6 mg/mL.

Mass Spectroscopy

8M urea in 200 mM phosphate pH 7.5 was added to solubilized crystals and protected with EZ-Link™ Maleimide-PEG2-Biotin in 20-fold molar excess at room temperature for two hours. Protein sample was run on a NuPage4-12% Bis-Tris gel (Invitrogen) and two gel fractions covering each gel band were de-stained and processed with iodoacetamide for alkylation, and further digested by trypsin without reduction and alkylation treatments. The digested peptide mixture was then concentrated and desalted using a C18 Zip-Tip. Mass spectroscopy analysis of disulfides was performed at Poochon Scientific LLC (Frederick, MD).

The desalted peptides were reconstituted in 26 µL of 0.1% formic acid. 2x12µl of peptides were analyzed by LC/MS/MS using a Thermo Scientific Q-Exactive hybrid Quadrupole- Orbitrap Mass Spectrometer and a Thermo Dionex UltiMate 3000 RSLCnano System. Samples were loaded onto a peptide trap cartridge at a flow rate of 5 µL/min and eluted onto a reversed-phase PicoFrit column (New Objective, Woburn, MA) using a linear gradient of acetonitrile (3-36%) in 0.1% formic acid for 60 min at a flow rate of 0.3 µL/min. Eluted peptides were ionized and sprayed into the mass spectrometer, using a Nanospray Flex Ion Source ES071 (Thermo) with spray voltage, 1.8 kV, capillary temperature, 250°C.

MS files were analyzed using the Thermo Proteome Discoverer 1.4.1 platform (Thermo Scientific, Bremen, Germany) for peptide identification and protein assembly. Database search against mouse protein sequences from NCBI website was performed using the SEQUEST and percolator algorithms through the Proteome Discoverer 1.4.1 platform. Cys-H (-1.008 Da), Cys-oxidation (+15.995 Da), oxidation of M (+15.995 Da) and deamidation Q/N (+0.98402 Da) were set as dynamic modifications. Minimum peptide length was 5 amino acids, precursor mass tolerance was 15 ppm, fragment mass tolerance was 0.05 Da and maximum false peptide discovery rate was 0.01. The estimation of relative abundance of protein was based on peptide spectrum match counts (PSM#). For identification of disulfide cross-linked peptides, the MS raw data file was analyzed using StavroX3.6.6 software (http://stavrox.com) using the following parameters: precursor precision 10.0 ppm, fragment ion precision 0.5 Da, lower mass limit 200.0 Da, upper mass limit 6000.0 Da, S/N ratio 2.0. Tryptic peptides resulted from trypsin digestion of the gel-band was analyzed by LC/MS/MS.

A minor band apparent in SDS PAGE was similarly analyzed using both trypsin and chymotryptic digests which confirmed that the protein was full length.

Mass Photometry

Dissolved crystals were diluted to 40 nM with PBS immediately before being run on an Interferometric Light Scattering Microscope (ISCAMS) which is able to quantify the mass of single biomolecules in solution with approximately 2% sequence mass accuracy, up to 19 kDa resolution, and 1 kDa precision with a detection limit of 50 kDa. The video file was processed with Discover MP [44]. The contrast values were converted to molecular masses using the instrument calibration function. The MP data was plotted as histograms or kernel density estimate (KDE) distributions. The distribution peaks were fit with Gaussian functions to obtain the average molecular mass of each distribution component.

Glutaraldehyde Crosslinking

9 µl of 0.1 mg/mL dissolved crystal or uncrystallized monomer in PBS was mixed with 1 µL 0.2% glutaraldehyde. 1 µL of 1M Tris-HCl, pH 8.0 was added with 3 µL of SDS-PAGE loading buffer

and incubated at 25 °C for twenty minutes. The samples were analyzed by SDS-PAGE. The gels were stained with AcquaStain Protein Gel Stain (Bulldog Bio).

Fluorescence

Fluorescence emission spectra were measured on a PTI Quantimaster fluorimeter (Photon Technology International, Trenton, NJ, USA) at 20 °C. 1nm slit widths were used with excitation at 280 mm and emissions spectra collected from 300-450nm. Protein solution was 0.1 mg/mL and buffer was 50mM sodium acetate, pH 5.0 with and without 7M GdnHCl.

Thermal stability

Turbidity was measured by OD_{600} scan of 0.2 mg/mL of dissolved crystals (tenfold dilution) or uncrystallized monomer in 50 mM sodium acetate, pH 5.0 at 60°C over a period of two hours using 1.5 mL disposable cuvettes in a Genesys 150 UV/VIS spectrophotometer (Thermo Scientific). Duplicate preparations were tested both freshly dissolved and after 24 hrs in solution. Since the dissolved crystals at this concentration contained 3% glycerol, the same amount was added to the uncrystallized protein.

Accession Number

PDB ID 7RJ0

Acknowledgements

This work was funded by the Intramural Program of the National Eye Institute. Diffraction data were collected at the Advanced Photon Source (APS), Argonne National Laboratory, Lemont, IL. We gratefully acknowledge the expert help of Drs Grzegorz Piszczek and Di Wu of the Biophysical Core NHLBI/NIH and Dr Jon Hao of Poochon Scientific.

Conflicts of Interest

None

Author Contributions

VS and GW designed experiments and analyzed data. VS performed experiments. GW and VS wrote the manuscript and prepared figures.

References

1. Chiti, F. & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75, 333-66.

2. Lee, S., Mahler, B., Toward, J., Jones, B., Wyatt, K., Dong, L., Wistow, G. & Wu, Z. (2010). A single destabilizing mutation (F9S) promotes concerted unfolding of an entire globular domain in gammaS-crystallin. *J Mol Biol* 399, 320-30.
3. Mahler, B., Doddapaneni, K., Kleckner, I., Yuan, C., Wistow, G. & Wu, Z. (2011). Characterization of a transient unfolding intermediate in a core mutant of gammaS-crystallin. *J Mol Biol* 405, 840-50.
4. Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6, 197-208.
5. Ferina, J. & Daggett, V. (2019). Visualizing Protein Folding and Unfolding. *J Mol Biol* 431, 1540-1564.
6. Bloemendal, H., De Jong, W., Jaenicke, R., Lubsen, N. H., Slingsby, C. & Tardieu, A. (2004). Ageing and vision: structure, stability and function of lens crystallins. *Prog Biophys Mol Biol* 86, 407-85.
7. Slingsby, C., Wistow, G. J. & Clark, A. R. (2013). Evolution of crystallins for a role in the vertebrate eye lens. *Protein Sci* 22, 367-80.
8. Rocha, M. A., Sprague-Piercy, M. A., Kwok, A. O., Roskamp, K. W. & Martin, R. W. (2021). Chemical Properties Determine Solubility and Stability in betagamma-Crystallins of the Eye Lens. *Chembiochem* 22, 1329-1346.
9. Slingsby, C. & Wistow, G. J. (2014). Functions of crystallins in and out of lens: Roles in elongated and post-mitotic cells. *Prog Biophys Mol Biol* 115, 52-67.
10. Zhao, H., Chen, Y., Rezabkova, L., Wu, Z., Wistow, G. & Schuck, P. (2014). Solution properties of gamma-crystallins: Hydration of fish and mammal gamma-crystallins. *Protein Sci* 23, 88-99.
11. Chen, Y., Zhao, H., Schuck, P. & Wistow, G. (2014). Solution properties of gamma-crystallins: compact structure and low frictional ratio are conserved properties of diverse gamma-crystallins. *Protein Sci* 23, 76-87.
12. Zhao, H., Chen, Y., Rezabkova, L., Wu, Z., Wistow, G. & Schuck, P. (2014). Solution properties of gamma-crystallins: hydration of fish and mammal gamma-crystallins. *Protein Sci* 23, 88-99.
13. Sagar, V., Chaturvedi, S. K., Schuck, P. & Wistow, G. (2017). Crystal Structure of Chicken gammaS-Crystallin Reveals Lattice Contacts with Implications for Function in the Lens and the Evolution of the betagamma-Crystallins. *Structure* 25, 1068-1078 e2.
14. Chaturvedi, S. K., Sagar, V., Zhao, H., Wistow, G. & Schuck, P. (2019). Measuring Ultra-Weak Protein Self-Association by Non-ideal Sedimentation Velocity. *J Am Chem Soc* 141, 2990-2996.
15. Wistow, G. (2012). The human crystallin gene families. *Hum Genomics* 6, 26.
16. Hejtmancik, J. F. & Smaoui, N. (2003). Molecular genetics of cataract. *Dev Ophthalmol* 37, 67-82.
17. Sandilands, A., Hutcheson, A. M., Long, H. A., Prescott, A. R., Vrensen, G., Loster, J., Klopp, N., Lutz, R. B., Graw, J., Masaki, S., Dobson, C. M., MacPhee, C. E. & Quinlan, R. A. (2002). Altered aggregation properties of mutant gamma-crystallins cause inherited cataract. *Embo J* 21, 6005-14.
18. Moreau, K. L. & King, J. A. (2012). Protein misfolding and aggregation in cataract disease and prospects for prevention. *Trends Mol Med* 18, 273-82.
19. Lampi, K. J., Wilmarth, P. A., Murray, M. R. & David, L. L. (2014). Lens beta-crystallins: the role of deamidation and related modifications in aging and cataract. *Prog Biophys Mol Biol* 115, 21-31.
20. Harding, J. J. & Dilley, K. J. (1976). Structural proteins of the mammalian lens: A review with emphasis on changes in development, aging and cataract. *Exp Eye Res* 22, 1-73.
21. Bron, A. J., Vrensen, G. F., Koretz, J., Maraini, G. & Harding, J. J. (2000). The ageing lens. *Ophthalmologica.* 214, 86-104.

22. Serebryany, E., Yu, S., Trauger, S. A., Budnik, B. & Shakhnovich, E. I. (2018). Dynamic disulfide exchange in a crystallin protein in the human eye lens promotes cataract-associated aggregation. *J Biol Chem* 293, 17997-18009.
23. Serebryany, E., Woodard, J. C., Adkar, B. V., Shabab, M., King, J. A. & Shakhnovich, E. I. (2016). An Internal Disulfide Locks a Misfolded Aggregation-prone Intermediate in Cataract-linked Mutants of Human γD-Crystallin. *Journal of Biological Chemistry* 291, 19172-19183.
24. Serebryany, E. & King, J. A. (2014). The betagamma-crystallins: native state stability and pathways to aggregation. *Prog Biophys Mol Biol* 115, 32-41.
25. Serebryany, E., Takata, T., Erickson, E., Schafheimer, N., Wang, Y. & King, J. A. (2016). Aggregation of Trp > Glu point mutants of human gamma-D crystallin provides a model for hereditary or UV-induced cataract. *Protein Sci* 25, 1115-28.
26. Wistow, G., Bernstein, S. L., Wyatt, M. K., Behal, A., Touchman, J. W., Bouffard, G., Smith, D. & Peterson, K. (2002). Expressed sequence tag analysis of adult human lens for the NEIBank Project: over 2000 non-redundant transcripts, novel genes and splice variants. *Mol Vis* 8, 171-84.
27. Wistow, G., Wyatt, K., David, L., Gao, C., Bateman, O., Bernstein, S., Tomarev, S., Segovia, L., Slingsby, C. & Vihtelic, T. (2005). gammaN-crystallin and the evolution of the betagamma-crystallin superfamily in vertebrates. *FEBS J.* 272, 2276-91.
28. Chen, Y., Sagar, V., Len, H. S., Peterson, K., Fan, J., Mishra, S., McMurtry, J., Wilmarth, P. A., David, L. L. & Wistow, G. (2016). gamma-Crystallins of the chicken lens: remnants of an ancient vertebrate gene family in birds. *FEBS J* 283, 1516-30.
29. Thorn, D. C., Grosas, A. B., Mabbitt, P. D., Ray, N. J., Jackson, C. J. & Carver, J. A. (2019). The Structure and Stability of the Disulfide-Linked gammaS-Crystallin Dimer Provide Insight into Oxidation Products Associated with Lens Cataract Formation. *J Mol Biol* 431, 483-497.
30. Bari, K. J., Dube, D., Sharma, S. & Chary, K. V. R. (2019). A Molecular Dynamics Perspective To Identify Precursors to Aggregation in Human gammaS-Crystallin Unravels the Mechanism of Childhood Cataracts. *J Phys Chem B* 123, 10384-10393.
31. Pike, A. C., Garman, E. F., Krojer, T., von Delft, F. & Carpenter, E. P. (2016). An overview of heavy-atom derivatization of protein crystals. *Acta Crystallogr D Struct Biol* 72, 303-18.
32. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.
33. Rudolph, R., Siebendritt, R., Nesslauer, G., Sharma, A. K. & Jaenicke, R. (1990). Folding of an all-á protein: independent domain folding in gamma II-crystallin from calf eye lens. *Procedings of the National Academy of Sciences USA* 87, 4625-4629.
34. Mayr, E. M., Jaenicke, R. & Glockshuber, R. (1997). The domains in gammaB-crystallin: identical fold-different stabilities. *J Mol Biol* 269, 260-9.
35. Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. (1995). 3D domain swapping: a mechanism for oligomer assembly. *Protein Science* 4, 2455-2468.
36. Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L. & Slingsby, C. (1990). X-ray analysis of betaB2-crystallin and evolution of oligomeric lens proteins. *Nature* 347, 776-780.
37. Van Montfort, R. L., Bateman, O. A., Lubsen, N. H. & Slingsby, C. (2003). Crystal structure of truncated human betaB1-crystallin. *Protein Sci* 12, 2606-12.
38. Xi, Z., Whitley, M. J. & Gronenborn, A. M. (2017). Human betaB2-Crystallin Forms a Face-en-Face Dimer in Solution: An Integrated NMR and SAXS Study. *Structure* 25, 496-505.
39. Schmidt, B., Ho, L. & Hogg, P. J. (2006). Allosteric disulfide bonds. *Biochemistry* 45, 7429-33.
40. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372, 774-97.

41. Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B. & Wistow, G. (1981). The molecular structure and stability of the eye lens: x-ray analysis of gamma-crystallin II. *Nature* 289, 771-7.
42. Hemmingsen, J. M., Gernert, K. M., Richardson, J. S. & Richardson, D. C. (1994). The tyrosine corner: a feature of most Greek key beta-barrel proteins. *Protein Science* 3, 1927-1937.
43. Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P. & Blundell, T. (1983). X-ray analysis of the eye lens protein gamma-II crystallin at 1.9 A resolution. *J Mol Biol* 170, 175-202.
44. Wu, D. & Piszczek, G. (2021). Standard protocol for mass photometry experiments. *Eur Biophys J* 50, 403-409.
45. Iadanza, M. G., Jackson, M. P., Hewitt, E. W., Ranson, N. A. & Radford, S. E. (2018). A new era for understanding amyloid structures and disease. *Nat Rev Mol Cell Biol* 19, 755-773.
46. Hyttinen, J. M., Amadio, M., Viiri, J., Pascale, A., Salminen, A. & Kaarniranta, K. (2014). Clearance of misfolded and aggregated proteins by aggrephagy and implications for aggregation diseases. *Ageing Res Rev* 18, 16-28.
47. Ciechanover, A. & Kwon, Y. T. (2015). Degradation of misfolded proteins in neurodegenerative diseases: therapeutic targets and strategies. *Exp Mol Med* 47, e147.
48. Zhao, L., Chen, X. J., Zhu, J., Xi, Y. B., Yang, X., Hu, L. D., Ouyang, H., Patel, S. H., Jin, X., Lin, D., Wu, F., Flagg, K., Cai, H., Li, G., Cao, G., Lin, Y., Chen, D., Wen, C., Chung, C., Wang, Y., Qiu, A., Yeh, E., Wang, W., Hu, X., Grob, S., Abagyan, R., Su, Z., Tjondro, H. C., Zhao, X. J., Luo, H., Hou, R., Jefferson, J., Perry, P., Gao, W., Kozak, I., Granet, D., Li, Y., Sun, X., Wang, J., Zhang, L., Liu, Y., Yan, Y. B. & Zhang, K. (2015). Lanosterol reverses protein aggregation in cataracts. *Nature* 523, 607-11.
49. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66, 213-21.

Table 1: Crystallographic statistics

Data collection	
Space group	$P2_12_12$
Cell dimensions *a*, *b*, *c* (Å)	61.09, 77.59, 154.05
α, β, γ (°)	90.0, 90.0, 90.0
Molecules per asymmetric unit	4
Wavelength (Å)	0.9793
Resolution (Å)	50–2.9
R_{sym} (last shell)	0.158 (0.550)
$I/\sigma I$ (last shell)	25 (3.3)
Completeness (%) (last shell)	99.3 (99.9)
Redundancy (last shell)	10.0 (10.6)
No. of reflections	16343
Refinement	
Resolution (Å)	48.0-2.9
R_{work}/R_{free} (%)	19.8/28.9
No. of protein residues in monomers A/B/C/D/	177/177/177/176
No. of non-protein atoms	
Solvent	69
Mean *B*-factors (Å^2)	89.9
RMSD	
Bond lengths (Å)	0.01
Bond angles (°)	1.29
MOLPROBITY	
Ramachandran outliers	1.3%
Ramachandran favored	87.8%
Clashscore	19.38%
Rotamer outliers	3.6%

Table 2

Side chain	Chi1(X1)	Chi2(X2)	Chi3(X3)	Distance (Angstrom)	Chi2'(X2')	Chi1'(X1')	Side Chain	Disulfide Strain Energy (kJ/mol)
C24 s2	-45.01	-74.41	-81.98	2.05	-81.87	-49.11	C24 s3	9.417402
C24 s4	-74.54	-38.93	-118.33	2.04	-38.93	-74.54	C24 s4	20.815723

Parameters for the C24-C24 disulfide bonds for s2/s3 and s4/s4.

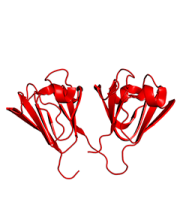
Table 3. RMSD comparison of N- and C-domains of s1-s4 with chicken γS-crystallin.

Domain	s1	s2	s3	s4
N: residues (4-84)	0.430	0.485	0.396	0.484
C: residues (93-177)	0.498	0.621	0.782	1.089

Table 4: Mass spectroscopy detection of disulfide bonds.

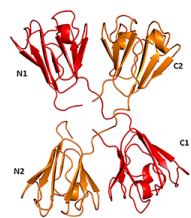
S-S	%
C24-C24	38.8
C22-C22	7.3
C22-C24	14.9
C22-C26	16.4
C24-C26	8.8
C26-C26	13.9

Mass spectroscopic analysis of tryptic peptides of gel-separated dimer slices, equivalent to the band shown in Fig 5C. Average of data for two slices. Analysis using StavroX3.6.6 shows evidence for intermolecular crosslinking of the C24 containing tryptic peptide with the most likely linkage between C24-C24 pairs.

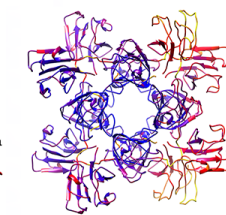


Monomer

Highly stable well-ordered



Twisted dimer



Asymmetric octamer

> disorder, aggregation