HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2017 November 13.

Published in final edited form as:

Science. 2017 August 11; 357(6351): 564–569. doi:10.1126/science.aal2674.

Eye patches: protein assembly of index-gradient squid lenses

J. Cai[1], **J.P. Townsend**[2], **T.C. Dodson**[1], **P.A. Heiney**[1], and **A.M. Sweeney**[1,*]

[1]University of Pennsylvania, Department of Physics & Astronomy.

[2]University of Pennsylvania, Department of Biochemistry and Biophysics.

Abstract

A parabolic relationship between lens radius and refractive index allows spherical lenses to avoid spherical aberration. We show that in squid, patchy colloidal physics resulted from an evolutionary radiation of globular S-crystallin proteins. Small angle x-ray scattering experiments on lens tissue show colloidal gels of S-crystallins at all radial positions. Sparse lens materials form via low-valence linkages between disordered loops protruding from the surface. The arms are polydisperse and bind via a set of hydrogen bonds between disordered side chains. Peripheral lens regions with low particle valence form stable, volume-spanning gels at low density, while central regions with higher average valence gel at higher densities. The proteins demonstrate an evolved set of linkers for self-assembly of nanoparticles into volumetric materials.

Maxwell showed that a “perfect medium” in a spherical lens would require a parabolic relationship between radial position and refractive index. This optical design has evolved in parallel in both fishes and decapod squid (1, 2), since in the light-limited ocean, this strategy optimizes sensitivity and acuity (Fig. 1A) (3). Here we describe the physical and material mechanisms underlying the evolution of this medium in squid.

In amorphous biological materials such as lenses, the concentration of material in aqueous cellular fluid determines the refractive index. The squid lens is cellular, so to a first approximation it might be possible to build a graded-index lens from protein solutions in an array of cells in which each cell has a higher concentration than the cell next to it. However, dissolved proteins would not likely result in a transparent lens. Most proteins interact attractively, such that aging of these solutions would lead to aggregation and light scattering. At some concentrations thermal motion can also cause light scattering (4, 5). Any aging is especially problematic since squid lens proteins do not undergo renewal: ribosomes and nuclei are strongly light scattering and are expelled from the developing cells (6).

“S-crystallins” are the set of proteins that constitute the refractive material in squid lens; they are all close variants of the glutathione s-transferase (GST) enzyme (7). This group of proteins has experienced episodes of strong positive (i.e., directional) evolutionary selection

*Correspondence to: alisonsw@physics.upenn.edu.

Supplementary materials:
Materials and methods
Figures S1–S4,
References 33–43.

relative to GST, suggesting that the many similar genetic copies in fact serve different functional roles (8). One of the most obvious but unexplained features of S-crystallins is a pair of unstructured peptide loops protruding from the folded, globular surface of the protein that bear little sequence similarity to other proteins (7, 8). These disordered sequences are encoded at the center of the linear amino acid sequence and the protein folds on either side of this central disordered region. Since S-crystallins are dimeric, the rotational symmetry of the dimer results in a neighboring pair of loops protruding from one end. The lengths of these loops and the surface charges of the proteins in the group vary (8).

The thermodynamic behavior of protein mixtures is typically rationalized using thermodynamic theories for isotropic particles (9, 10). These theories, however, cannot account for the absence of local density fluctuation throughout the lens that occurs with S-crystallins. We find that “patchy colloid” theory (11) can account for this apparent paradox of a gradient of protein volume fractions from 0–1 that escapes aging and local density fluctuation at all points in the gradient. This theory considers the interactions of colloids, or particles that are much larger than atoms, but small enough to undergo thermal collisions. The particles are considered “patchy” because they interact through hard-sphere repulsion plus attractive potentials that are localized to a few angularly small patches on the surface (11, 12). A surprising prediction of patchy colloid theory is that as particles’ average valence (the average number of patches per particle) approaches two, stable liquid-like materials form at arbitrarily low densities (11, 13). Conversely, as the average valence of the particles increases, the minimum density at which the system will gel can also be controlled (12, 13).

Results

Lens gradient

The relationship Maxwell described between radial position and refractive index is shown here for a spherical lens that focuses an aberration-free image for a single wavelength onto the retina (Fig. 1B, (1, 14)):

$$n(r)=\frac{n_c}{1+\left(\frac{n_c}{n_e}-1\right)r^2} \quad \text{Eq. 1}$$

Here, $n(r)$ is the index at the dimensionless, fractional radial position r, n_c is the index at the center of the lens, n_e is the index at the peripheral edge of the lens (16). In fish and squid lenses, n_c approaches the theoretical maximum for dry protein (~1.6), and n_e is close to that of the seawater in which the lens is immersed (~1.34) (14, 15).

We independently characterized the density gradient in squid lens and verified three premises: 1) that it is an increase in S-crystallin concentration alone, not a change in average amino acid refractive increment, that causes the observed relationship between radius and index; 2) that the shape of this relationship is consistent with the description in eq. 1; and 3) that the lower and upper bounds of the index range are 1.33 and 1.62, giving protein packing fraction (Φ) of 0.04 to ~1.0 (8,16–17). To partition the continuous density gradient of the lens into an experimentally tractable number of samples, we dissected tissue from individual

lenses into four concentric sections, with the outer edge of each subsample extending inward from fractions of roughly 100%, 80%, 60%, and 40% of the total lens radius (Fig. 1B). We refer to these concentric tissue samples as the 100%, 80%, 60% and 40% layers, and results from these samples are color-coded with red, orange, green, and blue, respectively, in the figures.

RNA sequencing

Given the age of extant S-crystallin studies and subsequent improvements in sequencing technologies, it was important to ensure we had captured the full range of S-crystallin genetic diversity expressed in the squid lens (17–19). Therefore, we sequenced transcriptomes of mature lens tissue from *Doryteuthis pealeii.* Our Illumina RNAseq resulted in 65.8 million paired-end reads for the 100% layer and 88.6 million paired-end reads for the 60% layer. After assembly, we found 53 unique S-crystallins transcripts. The predicted protein sequences are very similar to known S-crystallin sequences, with 52.0% ± 15.0% amino acid identity and 60.9% ± 12.9% similarity to Lops8, an archived S-crystallin from *L. opalescens* chosen for comparison (Genbank U19296.1) (18). In the 53 unique S-crystallins in the transcriptome, we identified 43 unique loop sequences that varied in length from 3 to 110 amino acids (Fig. 2A). 90% of the residues within the set of unique loops are algorithmically predicted to be disordered in solution, and are likely to encode an unstructured region of the protein polymer extending from the folded body of the homodimeric protein (8, 19–20).

Molecular weight characterization by gel electrophoresis

We used sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to characterize the protein molecular weight distribution as a function of lens radius. At all radial positions in the lens, more than 95% of the total protein was found in two broad molecular weight classes consistent with S-crystallin monomers, one with average molecular weight ~24.5 kDa and the other with average molecular weight ~26.5 kDa (Fig. 2B). There are also two minor bands in each sample, one at 32 and another at 36 kDa, consistent with S-crystallins from our transcriptome data with ~100 residue-long loop sequences (Fig. 2B). The molecular weight differences in SDS-PAGE are consistent with translations of our lens transcriptome, such that differences in S-crystallin molecular weight with lens radius are likely due to different loop composition with lens radius.

We performed a deconvolution of our SDS-PAGE data by modeling each migration pattern as a sum of Gaussian contributions from individual proteins with molecular weights encoded by full-length sequences obtained from RNA-seq (16). This analysis shows that although the ratio of the two major peaks centered at 24.5 and 26.5 kDa (due to the presence of short and long loops, respectively) shifts continuously toward larger proteins as a function of increasing radius, the underlying mixture of protein sequences associated with this shift is complex, not the result of a binary mixture of a few long- and short-loop sequences (Fig. 2B). The third peak on the gel, due to proteins with ~100 amino-acid-long loops, is a minor constituent of the periphery, reaches a maximum of 5% total protein at a relative radial position of 0.36, and is also present as 5% of total protein in the core. Proteins with very

short loops are highly abundant in the core (34% of total protein), but a minor constituent in the peripheral layer (3% of total protein).

Small-angle x-ray scattering (SAXS)

We used SAXS to understand the spatial structure of S-crystallins within the squid lens. The scattering intensity as a function of wave vector ($I(q)$) was found to be very similar for each layer of the lens, even though the packing fraction of the layers systematically changes from $\Phi = 0.05$ to $\Phi \approx 1.0$ (Fig. 3A). In all samples, at small wave vector $q = 4\pi \sin(\theta)/\lambda$, where θ is the half angle between the x-ray beam and the detector, the scattering intensity for all layers decreases with increasing q.. There were three major features also observed in scattering from all radial position within the lens: a broad, low peak between 0.01 $Å^{-1} < q < 0.03$ $Å^{-1}$, a shoulder around q=0.15 $Å^{-1}$, and a second peak near q=0.21 $Å^{-1}$. The shoulder near q=0.15 $Å^{-1}$ has a position and shape similar to scattering from a solution of 5 nm spheres. The peak at q= 0.21 $Å^{-1}$ corresponds to the real-space distance between the monomers within an S-crystallin dimer of ~3 nm.

Dilution and scattering form factor

To obtain a form factor, or scattering from an isolated protein, we initially attempted to dilute lens proteins in buffer. This procedure resulted in transparent suspensions containing >95% of the lens material when using tissue from the periphery of the lens. When treated in the same manner, samples from the core resulted in an insoluble pellet comprising >95% of the lens material. Tissue from intermediate radial positions resulted in a systematic gradient between these two extremes (Figs. S1, 4A). SAXS characterization of the supernatants from each radial position revealed structures consistent with folded protein at high q, but with a relationship of $S(q) \propto q^{-1.5}$ between 9e-3≥q≥4e-3, the minimum q technically accessible. (Figs. S2, S3). This result was independent of the salt content of the buffer, and the concentration of protein in the fluid in a concentration range from ~10 mg/mL (Φ = 1.4e-2) to ~0.05 mg/mL (Φ = 1.1e-4), the minimum at which SAXS data were meaningful (Fig. S4). A structural fit to these SAXS data using the algorithm AMBIMETER produced a low-ambiguity result that these solutions contain pairwise-linked chains of S-crystallin proteins (21). Patchy colloid theory gives insight to this behavior: in patchy systems, dilution will not cause a transition from glassy to isolated particles, as is the case for isotropic particles (9, 10), but instead will cause a transition from an equilibrium gel to an equilibrium percolating/cluster fluid (13).

This particle assembly upon dilution meant that it was not possible to measure a form factor in the usual manner. Therefore, we calculated a form factor of isolated S-crystallin dimers using atomic positions taken from S-crystallin homology models (16). The predicted form factors of all S-crystallin homodimers are similar: they are flat at small values of $q < 0.05$ $Å^{-1}$, then decrease to a minimum near $q = 0.17$ $Å^{-1}$, with a minor peak at $q \approx 0.21$ $Å^{-1}$ (Fig. 3A).

Structure Factor, S(q)

The “structure factor” is a reciprocal-space representation of the interactions between defined particles in a sample, found by dividing the overall scattering of a sample by the

form factor. The Fourier transform of the structure factor then becomes the pair-distribution function of the particles defined by the form factor. We calculated the structure factor of the lens tissue using the total SAXS scattering intensity and the average of the calculated form factors from individual S-crystallin dimers. This structure factor was also found to share a few prominent features for all radial positions of the lens: a strong peak in the structure factor at q =0.15 Å^{-1}, a trough at 0.03 Å^{-1}< q < 0.1 Å^{-1}, and at q < 0.01 Å^{-1} a relation $\Delta(\log I)/\Delta(\log q)$ of approximately -2.3 (Fig. 3B).

The most prominent peak in the structure factors at $q \approx 0.15$ Å^{-1} corresponds to the real-space size of the S-crystallin dimer. This suggests that the structure factor of all regions of the lens is dominated by pair-wise dimer-dimer interactions. This peak shifts toward higher q toward the center of the lens, meaning that the real-space pairwise distance decreases from the periphery to the core of the lens. Measurements on samples from different lens radii also shared a minimum in structure factor at $q \cong 0.05$ Å^{-1} that increased with decreasing lens radius. This relationship is similar to previous numerical work that found that this peak in $S(q)$ indicating pair-wise interactions shifted to the right with increasing numbers of contacts between proteins (coordination number or $\langle M \rangle$) (Fig. 3B, inset) (22).

The broad feature in structure factor between 0.01 Å^{-1} < q < 0.03 Å^{-1} could be due to a few, possibly interacting, phenomena. First, there could be local swelling of the gel due to weak van der Waals, electrostatic, hydrophobic interactions, or water migration within the lens (23). This feature could also indicate complex, multiparticle nodes within the protein network, consistent with those observed in simulations (22). Samples from all regions in the lens also had similar structure factors in the region 0.006 Å^{-1} < q < 0.01 Å^{-1}, where the signal exhibited a slope of slightly more than -2 in a log-log plot, consistent with expectations for a percolating network (24).

Dilution and characterization of network coordination number

We observed that when tissue from different radial positions was diluted to a volume fraction of 0.01 then centrifuged, two materials formed: a protein fluid and a pellet (Fig. S1). As described above, the sparse phase is a percolating protein fluid (Figs. S2–S4). Both the relative amount of protein in and the density of the pellet increased with decreasing lens radius (Fig. 4A). In samples taken from the periphery of the lens, the pellet was a soft, volume-spanning translucent gel, while the core of the lens formed a white powder (Fig. S1). Since all S-crystallins are extremely hydrophilic (mean grand average of hydropathy index per sequence of -5.61±0.07), this variation in pellet formation is due to different protein-protein interactions at different radial positions of the lens, not differences in protein-water interactions (25). Except for at the extreme peripheral edge of the lens, the pellet that formed was always less dense than the lens tissue that went into the sample, even after accounting for pore space of volume fraction 0.35 in the powder pellets. This result suggests that a phase or state transition of S-crystallins occurs upon dilution of the native tissue (Figs. 4A–C, S1).

This phase or state transition upon dilution allowed us explore the intrinsic average coordination number, $\langle M \rangle$, of the particles as a function of lens radius (11). Given the high sequence polydispersity present at every radial position in the lens (Fig. 2), a polydisperse

set of effective protein-protein bond temperatures must also be present. If this set of effective bond temperatures spans the spinodal line for the system, the diagram predicts that upon dilution the proteins will separate into a sparse phase containing the “high temperature” (low bond energy) particles located above the spinodal line, and a dense material containing the “low temperature” (high bond energy) particles (Fig. 4B) (11). The “low temperature” particles, when relocated to the unstable region of the phase space under the spinodal line, must then re-equilibrate to a density dictated by the average valence of this subset of the original total protein mixture (Fig. 4B–C). This phenomenon allowed us to estimate $<M>$ for the pellet material given its density. We used the measured particle density in each material, the inferred $<M>$ of the pellet, and the fact that in the spare chain the particles are in chains with $<M>=2$ to calculate $<M>$ as a function of radial position (Fig. 4A). (Fig. 4A). We found that $<M>$ in the periphery of the lens is slightly greater than two, and increased toward the center of the lens to a maximum near six at the core (Fig. 4A).

Monte Carlo Analysis

We used the algorithm DAMMIF, a Monte Carlo approach to finding a coarse-grain realspace structure given SAXS data (26). For the outermost layer of the lens, this result shows a network of 3–5 nm diameter spheres that are connected by pairs of “arms” leaving the body of an individual sphere (Fig. 1C). Most of the ~5 nm spheres show a rotational symmetry consistent with dimers of two structurally similar monomers. The distance bridged by these arms is typically ~40 Å, and the arms leave individual ~5 nm spheres at angles ranging from 60° to 120° with respect to each other. The average nearest-neighbor distance between the centers of two bridged proteins was 75 Å in the 100% layer and 60 Å in the 80% layer. These predictions are consistent with sequencing data and homology models that S-crystallins have extended loops protruding from the two pairs of helices forming the core of the folded structure (Figs. 2B, 3D, 5), with the loops linking the bodies of the proteins into a network of spherical particles.

Because DAMMIF is a Monte Carlo technique, each calculation generates a likely but not unique structural fragment given a single SAXS measurement. However, the volume-spanning material in the lens is much larger than the ~300-nm real-space span accessed by SAXS. Therefore, we generated an ensemble of 100 DAMMIF outputs for each layer of the lens, and treated these as ensembles of fragments of the larger protein network forming the lens.

This approach generated an independent estimate of $<M>$ as a function of lens radius and showed how $<M>$ emerges from S-crystallin secondary and tertiary structure. In the ensembles from the 100% and 80% layers, we observed 5-nm proteins connected by thin bridges that vary in length from 50 Å to 400 Å, to a minimum of one and a maximum of four other dimers (Fig. 3D). In the 100% layer, the majority of particles have two contacts, with the pair of unstructured loops from an S-crystallin dimer forming contacts with loops from other dimers. A few particles have a third contact resulting from loop-body interaction with another dimer. The 80% layer had a greater proportion of $M = 3$ nodes in the network than the 100% layer. In the 80% layer, a second form of $M=3$ interactions appears: loops of length ~100 amino acids bifurcate to form a Y-shaped structure. This Y-shaped structure

results in two points of contact within a single loop, each capable of contacting the end(s) of another loop (Fig. 3D).

These structural ensembles showed $<M>$ = 2.19 for the 100% layer, and $<M>$ = 2.45 for the 80% layer, and a permutation test showed the distribution of M to be different between the two ensembles. The dilution experiment found $<M>$ = 2.01–2.22 for the 100% layer and $<M>$ = 2.22–3.11 for the 80% layer (Fig. 4A), so two independent estimates of $<M>$ agree.

At samples from lens radii of 60% and 40%, the predicted structures were dense, with contacts between particle surfaces occupying much of the surface area. The 60% layer prediction showed 5 nm-thick ramifying lamellae, consistent with interconnected sheets of S-crystallin dimers (Fig. 3D). In the 40% layer prediction, protein completely filled the volume, consistent with a solid containing little or no bulk water (Fig. 3D). The dilution analysis described above found that $<M>$ =3.1–3.9 for the 60% layer, and $<M>$ = 3.9–5.8 for the 40% layer (Fig. 4A), qualitatively consistent with these structural predictions.

Network-of-spheres simulations

As an aid to visualization, we simulated branching chains of particles, generating non-equilibrium structures with qualitative similarity to those in our SAXS data, and calculated the resulting structure factors. (16). We added spheres at arbitrary angles to a growing chain of spheres, with the addition of two spheres causing a bifurcation in the chain (Fig. 3E). We chose bifurcation frequencies so that average particle valences were similar to those in the lens. For all simulated $<M>$, there was a peak at q = 0.15 $Å^{-1}$ corresponding to the particle-particle nearest-neighbor distance, a minimum in structure factor between $0.03<q<0.1$ $Å^{-1}$, and log-log slope of -2.3 for $q < 0.07$ $Å^{-1}$.

These simulations also show the origin of the local minimum in lens structure factor around q=0.10 $Å^{-1}$. There is systematically less spatial fluctuation at this q-range when the particles are more closely packed at higher $<M>$; their positions are more constrained with respect to each other, causing in a decrease in density fluctuation. The depth of the minimum at q=0.10 $Å^{-1}$ in the experimental data is deeper than in the simulation for radial positions less than 60%, further supporting the result that $<M>$ at these radial positions is likely greater than three.

Molecular details of protein-protein interactions

To study the atomic origin of M=2 interactions in the lens material, we used an explicit-water molecular dynamics (MD) approach using the coarse-grained structure of the 100% layer as a starting point (16). When we positioned two long-loop S-crystallins at the DAMMIF-predicted distance and relative orientation, several hydrogen bonds formed between the two disordered loops, forming a stable bridging structure (Fig. 5A). The number of hydrogen bonds increased from zero to a maximum six, with the resulting structure resembling a metastable amyloid interaction (Fig. 5B).

We used this simulation to estimate the bond energies resulting from these loop-loop contacts. We characterized protein-protein interactions with a dimensionless temperature, T, defined as the thermal energy of the system divided by the bonding energy of a particular

interaction. A previous study measured the Gibbs free energy of hydrogen bonding between hydrated strand-like peptides to average 2.6 kcal/H-bond/mol (27). For links consisting of 3–8 H-bonds at 10°C, T=0.08–0.03. This estimate of T applies to the 100% and 80% radius region of the lens where linkages between proteins are primarily between disordered loops.

At radial positions less than 60%, proteins form ramifying lamellae and then a proteinaceous material with no bulk water. These structures are consistent with patchy systems in which patches have become so large and/or numerous that their interactions approach those of isotropic, spherical particles (Figs. 3D, 4D). Here, the energies of protein-protein binding are likely similar to those involved in protein docking, in the range of 10–20 kcal/bond/mol, so the lower bound for T for these interactions is ~0.01 (28).

Discussion

Our data show that at all radial positions and every possible density, structure in the intact squid lens is dominated by attractive interactions between S-crystallin dimers, a result that is inconsistent with isotropic potentials, since these systems will undergo liquid-liquid phase separation and opacification at some point (29).

From a materials perspective, this lens could in principle be built from a single mixture of particles with M slightly greater than two, since any system to the right of the spinodal line for $<M>$=2.1 in the patchy colloid phase diagram may form a volumetric material with low density variation. However, our data show that $<M>$ changes monotonically with lens radius, from a minimum near two in the periphery to a maximum near six at the core of the lens.

There is a biological rationale for the lens evolving an array of $<M>$ rather than simply using one mixture with $<M>\approx$2.1 at many different densities. New lens cells must be filled gradually with protein through the process of mRNAs and amino acids diffusing to large ribosomes, with full-length proteins then diffusing away from the ribosome. In any cell whose proteome-determined $<M>$ allowed for gelation at a density lower than the appropriate endpoint for a given radial position, protein synthesis would be arrested prematurely. Patchy colloidal physics allows for a cellular mechanism by which lens cells gel near the optically appropriate endpoint density for a given radial position. The cellular transcriptome systematically titrates S-crystallin isoforms as $<M>$ changes as a function of lens radius. We hypothesize that a cell then passively ceases protein synthesis at the appropriate density, since upon gelation, protein synthesis would simply stop due to inability of large polymers (mRNAs and newly transcribed protein) to diffuse through the material.

Given that in principle one protein sequence could encode one valence in the lens, requiring a minimum of six S-crystallin sequences, there were a large number of unique S-crystallin sequences encoded by the genome, and the loops encoded by these sequences that form the linkages between proteins also had high chemical polydispersity. We speculate that this high degree of polydispersity in loop sequence and subsequent linker interaction dynamics may both avoid kinetic traps during nucleation and reduce the entropic cost of forming a volume-spanning network (30). This set of linker sequences generated by evolution may then help

inform experimental attempts to exploit the self-assembly principles revealed by patchy particle theory.

We observed that the difference between the original density of the lens and the density of the pellet formed from low effective temperature bonds increased from near zero at the periphery to a factor of two at the core (Fig. 4A–C). At the core of the lens, the pellet after a dilution-induced phase/state transition had a packing fraction near 0.5, indicating $<M>$ for the system near six, or square-lattice-like packing, while the intact lens core has little or no bulk water and a packing fraction approaching 1.0 (consistent with M=12). We speculate that in the squid embryo, these original, most central cells in the lens structure initially gel near the lower square-lattice packing fraction of 0.5, consistent with constituent proteins of $<M>\approx 6$. Then, as new cells are added to the organ periphery during lens growth, the newer material will be both more compliant due to a lower coordination number, and have a higher charge density due to the increasingly positive surface charge of the proteins in the peripheral layers (8). Both of these factors will tend to cause water to migrate outward from the lens core as the structure grows, systematically increasing the density of central regions of the lens with the animal's age and size, and potentially resulting in a near-dry lens core in the mature organ.

Our data also describe a gradual "inside-out" mechanism for evolving a gradient index lens from a single protein fold. In this view, an initial protein closely related to enzymatic glutathione s-transferase duplicated for lens expression and formed a high-density colloidal gel through near-isotropic interactions. Selection pressure for a gradient index then resulted in proteins with less-isotropic interactions and the resulting lamellar structures observed in extant lenses, consistent with patchy colloidal systems of intermediate average valence (31). Further selection on S-crystallins resulted in an unstructured loop encoded in a novel exon that was able to undergo flexible, non-specific hydrogen-bonding interactions with other similar loops. These loops, coupled to the increase in positive charge on the folded, spherical surfaces of the protein represent an evolutionary innovation of a protein particle able to enforce a coordinate number of M=2. With this loop-binding innovation, squid lenses were then able to exploit the entire physics of the patchy colloid phase diagram, and the thermodynamically stable materials with low density fluctuation that result from this physics. It is also possible that this patchy colloidal perspective could provide insights into still poorly understood aspects of vertebrate lens biology. In particular, the polydisperse nature of alpha crystallins in the low-density regions of the vertebrate lens have been a puzzle; even in solution, they appear to interact via "tentacles" in ways that are hard to specifically characterize (32). It is possible that they are also acting as low-valence patchy colloidal gels in vivo, as is the case for squid S-crystallins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to R. Kamien and E. Eiser for useful discussions, and to D. Fox for assistance with data visualization. S. Johnsen's comments improved the manuscript. We are also grateful to an anonymous reviewer

whose thoughtful questions improved the work. Financial support was provided by the National Science Foundation MRSEC DMR11-20901 to PAH, by the Packard Foundation Fellowship for Science and Engineering, Sloan Foundation, NSF-1351935, Kaufman Foundation, and University of Pennsylvania to AMS, and by the DoD NDSEG graduate fellowship program to TCD. SAXS data are archived at SASBDB accession numbers SASDCQ5–U5; and RNAseq data are archived at GenBank under SRR5528268–9.

References

1. Maxwell JC. Solutions of problems. Camb. Dublin Math J. 1854; 8:188.
2. Packard A. Cephalopods and fish: The limits of convergence. Biol. Rev. Biol. Proc. Cambridge Philos. Soc. 1972; 47:241–307.
3. Land, MF., Nilsson, DE. Animal eyes. Oxford University Press; 2012.
4. Malfois M, Bonneté F, Belloni L, Tardieu A. A model of attractive interactions to account for fluid–fluid phase separation of protein solutions. J. Chem. Phys. 1996; 105:3290–3300.
5. Benedek GB. Theory of transparency of the eye. Appl. Opt. 1971; 10:459–473. [PubMed: 20094474]
6. West JA, Sivak JG, Doughty MJ. Microscopical evaluation of the crystalline lens of the squid (*Loligo opalescens*) during embryonic development. Exp. Eye Res. 1995; 60:19–35. [PubMed: 7720802]
7. Tomarev SI, Zinovieva RD, Piatigorsky J. Characterization of squid crystallin genes. Comparison with mammalian glutathione s-transferase genes. J. Biol. Chem. 1992; 267:8604–8612. [PubMed: 1373730]
8. Sweeney AM, Des Marais DL, Ban YEA, Johnsen S. Evolution of graded refractive index in squid lenses. J. R. Soc., Interface. 2007; 4:685–698. [PubMed: 17293312]
9. Gibaud T, Najet M, Oberdisse J, Lindner P, Pedersen JS, Oliveira CL, Stradner A, Schurtenberger P. New routes to food gels and glasses. Faraday discussions. 2012; 158(1):267–284.
10. Cardinaux F, Zaccarelli E, Stradner A, Bucciarelli S, Farago B, Egelhaaf SU, Sciortino F, Schurtenberger P. Cluster-driven dynamical arrest in concentrated lysozyme solutions. The Journal of Physical Chemistry B. 2011; 115(22):7227–7237. [PubMed: 21528887]
11. Bianchi E, Largo J, Tartaglia P, Zaccarelli E, Sciortino F. Phase diagram of patchy colloids: Towards empty liquids. Phys. Rev. Lett. 2006; 97:168301. [PubMed: 17155440]
12. Gögelein C, Nägele G, Tuinier R, Gibaud T, Stradner A, Schurtenberger P. A simple patchy colloid model for the phase behavior of lysozyme dispersions. The Journal of chemical physics. 2008; 129(8):08B615.
13. Sciortino F, Zaccarelli E. Reversible gels of patchy particles. Current Opinion in Solid State and Materials Science. 2011; 15(6):246–253.
14. Jagger WS. The optics of the spherical fish lens. Vision. Res. 1992; 32:1271–1284. [PubMed: 1455702]
15. Sivak JG, West JA, Campbell MC. Growth and optical development of the ocular lens of the squid (*Sepioteuthis lessoniana*). Vision. Res. 1994; 34:2177–2187. [PubMed: 7941414]
16. Materials and methods are available online as supplementary materials
17. Tomarev SI, Zinovieva RD. Squid major lens polypeptides are homologous to glutathione s-transferases subunits. Nature. 1988; 336:86–88. [PubMed: 3185725]
18. Tomarev SI, Zinovieva RD, Guo K, Piatigorsky J. Squid glutathione s-transferase. Relationships with other glutathione s-transferases and s-crystallins of cephalopods. J. Biol. Chem. 1993; 268:4532–4542.
19. Tomarev SI, Chung S, Piatigorsky J. Glutathione s-transferase and s-crystallins of cephalopods: Evolution from active enzyme to lens-refractive proteins. J. Mol. Evol. 1995; 41:1048–1056. [PubMed: 8587103]
20. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. Nucleic acids research. 2008; 36(suppl 2):W197–W201. [PubMed: 18463136]
21. Petoukhov MV, Svergun DI. Ambiguity assessment of small-angle scattering curves from monodisperse systems. Acta Crystallographica Section D: Biological Crystallography. 2015; 71(5):1051–1058.

22. Russo J, Tartaglia P, Sciortino F. Reversible gels of patchy particles: Role of the valence. J. Chem. Phys. 2009; 131:014504. [PubMed: 19586107]
23. Shibayama M. Spatial inhomogeneity and dynamic fluctuations of polymer gels. Macromol. Chem. Phys. 1998; 199:1–30.
24. Stauffer, D., Aharony, A. Introduction to percolation theory. CRC Press; 1994.
25. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. Journal of molecular biology. 1982; 157(1):105–132. [PubMed: 7108955]
26. Franke D, Svergun DI. Dammif, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Crystallogr. 2009; 42:342–346. [PubMed: 27630371]
27. Eildal J, Hultqvist G, Balle T, Stuhr-Hansen N, Padrah S, Gianni S, Strømgaard K, Jemth P. Probing the role of backbone hydrogen bonds in protein–peptide interactions by amide-to-ester mutations. Journal of the American Chemical Society. 2013; 135(35):12998–13007. [PubMed: 23705582]
28. Janin J. Principles of protein-protein recognition from structure to thermodynamics. Biochimie. 1995; 77(7):497–505. [PubMed: 8589061]
29. Camp PJ. Phase diagrams of hard spheres with algebraic attractive interactions. Phys. Rev. E. 2003; 67:011503.
30. Jacobs WM, Reinhardt A, Frenkel D. Communication: Theoretical prediction of free-energy landscapes for complex self-assembly. J. Chem. Phys. 2015; 142:021101. [PubMed: 25591330]
31. Bianchi E, Capone B, Kahl G, Likos CN. Soft-patchy nanoparticles: Modeling and self-organization. Faraday Discuss. Chem. Soc. 2015; 181:123–138.
32. Vanhoudt J, Abgar S, Aerts T, Clauwaert J. A small-angle x-ray solution scattering study of bovine α-crystallin. Eur. J. Biochem. 2000; 267:3848–3858. [PubMed: 10849004]
33. Goldschmidt L, Teng PK, Riek R, McMeekin TL, Wilensky M, Groves ML. Refractive indices of proteins in relation to amino acid composition and specific volume. Biochem. Biophys. Res. Commun. 1962; 7:151–156.
34. Zhao H, Brown PH, Schuck P. On the distribution of protein refractive index increments. Biophys. J. 2011; 100:2309–2317. [PubMed: 21539801]
35. Putnam DK, Lowe EW, Meiler J. Reconstruction of saxs profiles from protein structures. Comput. Struct. Biotechnol. J. 2013; 8:1–12.
36. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput. 2008; 4:435–447. [PubMed: 26620784]
37. Case, DA., Betz, RM., Botello-Smith, W., Cerutti, DS., Cheatham, TE., III, et al. AMBER 2016. University of California; San Francisco: 2016.
38. Arnold K, Bordoli L, Kopp J, Schwede T. The swiss-model workspace: A web-based environment for protein structure homology modelling. Bioinformatics. 2006; 22:195–201. [PubMed: 16301204]
39. Haas BJ, et al. De novo transcript sequence reconstruction from rna-seq using the trinity platform for reference generation and analysis. Nat. Protocols. 2013; 8:1494–1512. [PubMed: 23845962]
40. Eddy SR. Profile hidden markov models. Bioinformatics. 1998; 14:755–763. [PubMed: 9918945]
41. Ji X, et al. Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens s-crystallins of cephalopods. Biochemistry. 1995; 34:5317–5328. [PubMed: 7727393]
42. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. Nucleic acids research. 2008; 36(suppl 2):W197–W201. [PubMed: 18463136]
43. Goldschmidt L, Teng PK, Riek R, Eisenberg D. ZipperDB: Predictions of Fibril-forming Segments within Proteins Identified by the 3D Profile Method (from the UCLAD-OE Institute for Genomics and Proteomics).

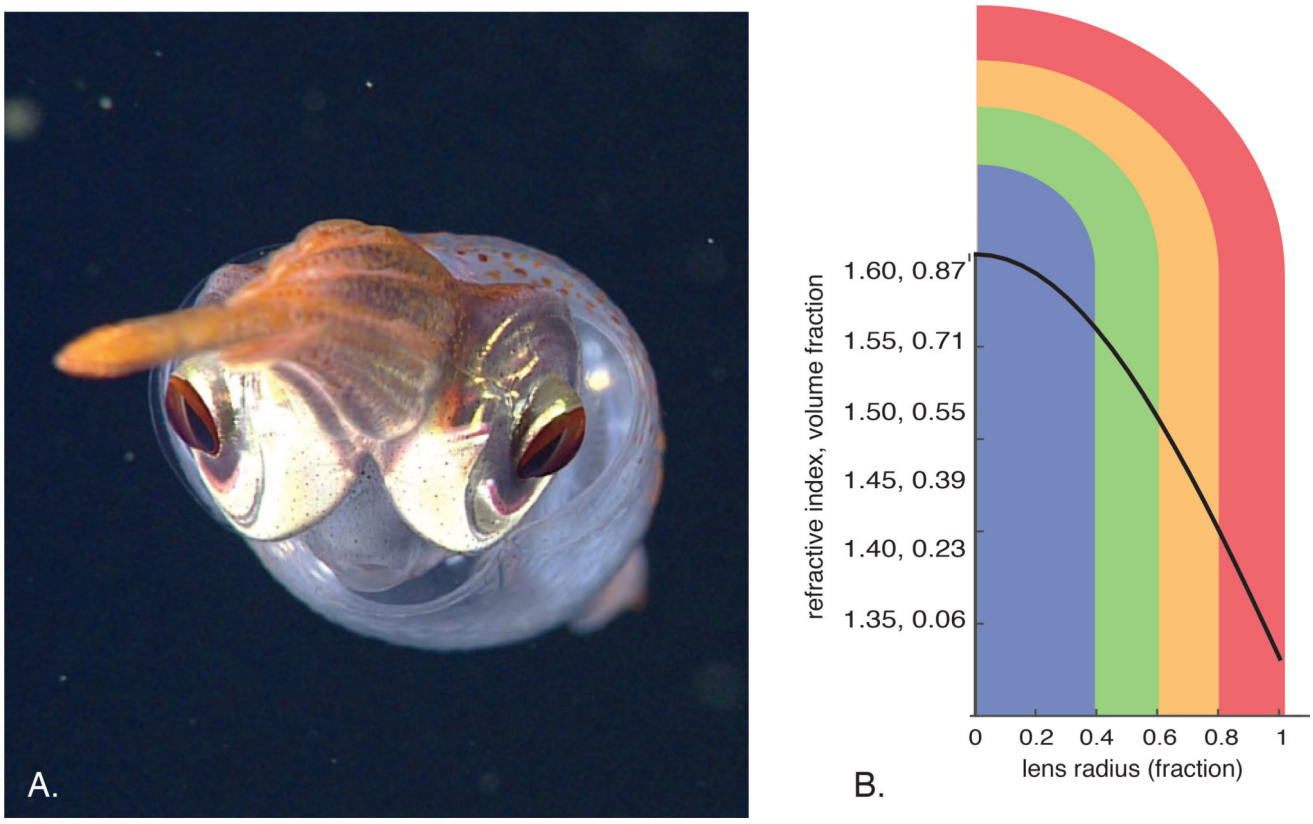

Figure 1.

Relationships between lens radius, refractive index, protein structure and network structure. **A)** In situ photograph of a decapod squid showing spherical, gradient index lenses. **B)** Lens radius vs. refractive index, with an overlay showing the relationship between lens radius and the color-coding used in this work to identify discretized tissue samples (blue for r<40%, green for 40%<r<60%, orange for radii 60%<r<80% and red for r>80%.

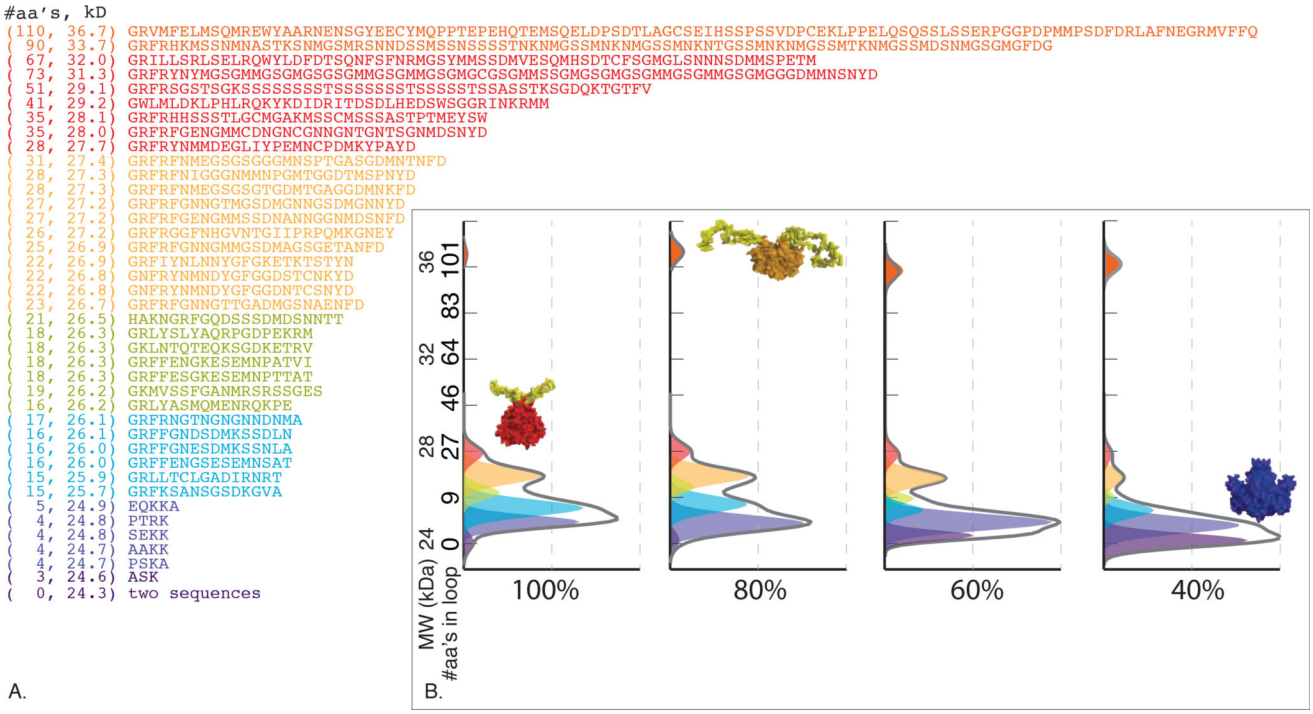

Figure 2.

"Loop" sequences and their estimated relative abundance as a function of lens radius. **A)** Unique "loop"-encoding sequences as determined by lens RNA sequencing, and the relative abundance of proteins of different molecular weights. Parentheses show number of amino acids in the sequence and its molecular weight in kilodaltons. Sequence text is colored to correspond to peak decomposition in (**B**). **B)** SDS-PAGE of lens samples taken from different radial positions of the lens, and an estimated abundance of components of different loop molecular weights. The grey curve shows both the raw SDS-PAGE staining density and the sum of the curves in the fit (the two are similar to within the width of the line). The colored regions show the estimated abundance of a molecular weight component within a given sample. Images of protein show the predicted homology model structure of an S-crystallin in a given molecular weight class (long loops in red, short loops in blue, and 100-aa loops in orange).

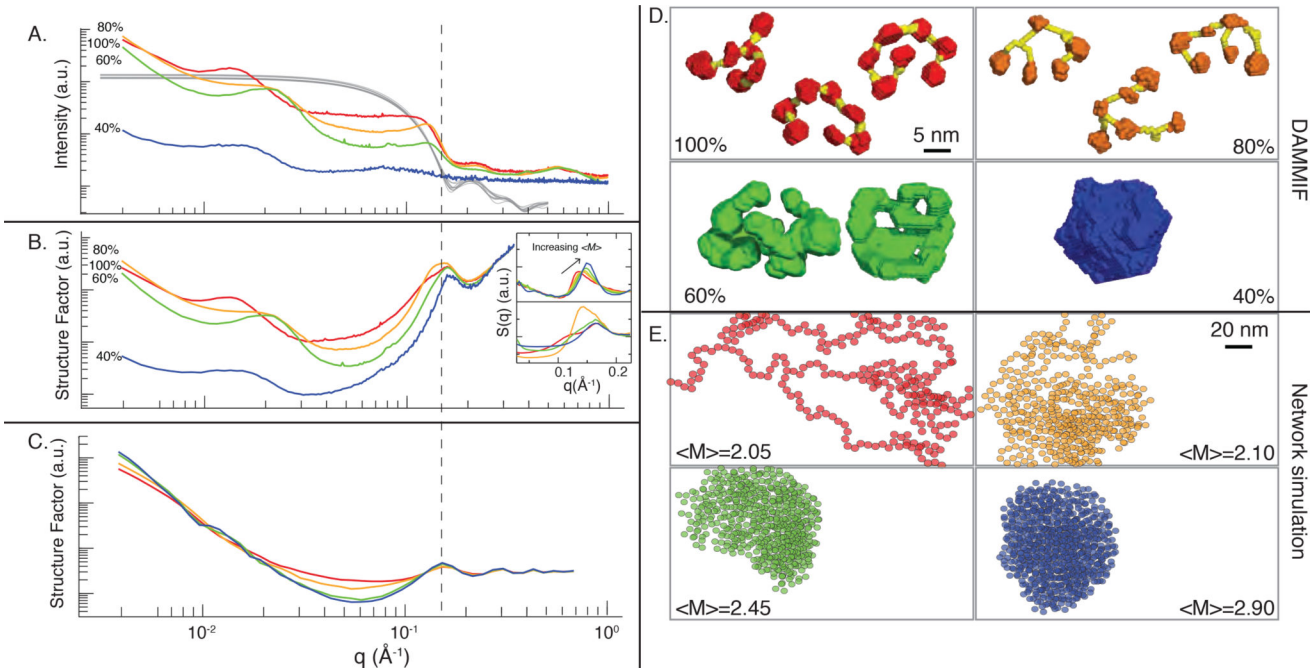

Figure 3.

SAXS, DAMMIF structures, and modeled networks of particles. <*M*> is the average protein coordination number. **A)** X-ray scattering intensity as a function of *q*, (I(*q*)). Measurements from different radial portions of the lens are indicated by color. Gray traces show the calculated form factor of isolated S-crystallin dimers. **B)** Structure factors of different concentric regions of the lens, calculated using *I(q)* and the calculated form factor shown in (**A**). **Inset of (B)**: Peak position associated with particle pair-wise interactions in numerical simulation (top panel, data from [*22*])) and experimentally measured in this study (bottom panel). The peak associated with pairwise interactions moves toward lower *q* as <*M*> increases when *Φ* is held constant. **C)** Structure factors of simulated branching networks of particles with coordinate numbers predicted to exist in the squid lens. **D)** Sample of DAMMIF predictions of lens material structure, calculated using the SAXS data in (**A**), coordination number. Scale bar applies to all panels of (**D**). Individual S-crystallins are colored according to radial position while predicted links between S-crystallins are colored yellow. **E)** Sample of structures from network simulations, whose structure factors are shown in (**C**). Scale bar applies to all panels of (**E**). 10,000 particles were included in the calculation of structure factor; a few hundred particles from each simulation are shown here, in a box of the same volume for each simulation.

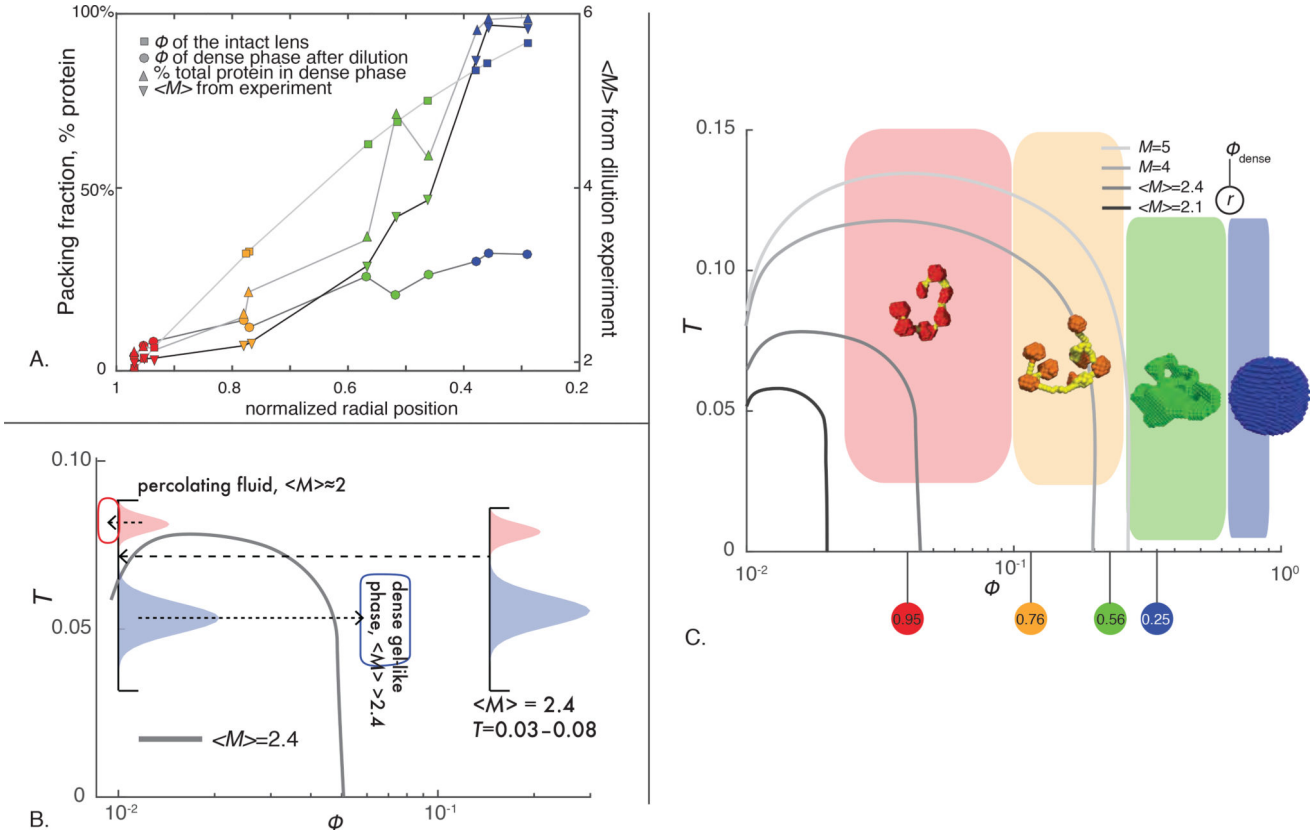

Figure 4.

Experimental data mapped to patchy colloid theory **A)** Volume fraction (Φ) of the intact lens, of the dense material after dilution, percent total protein in the dense state, and calculated $<M>$ resulting from this experiment as a function of lens radius. Each symbol represents one measurement and three lenses were sampled for each curve. **B)** Schematic showing how a polydisperse set of effective bond temperatures in a patchy colloidal system, when diluted, will phase separate. Vertical line indicates the likely range of effective bond temperatures (T) in a given mixture in the lens, and the filled red and blue shapes indicate a possible set of “high” and "low" effective temperature bonds in the system. Arrows indicate how “high temperature” particles in the system will re-equilibrate into a percolating fluid, while “low temperature” particles are in an unstable part of the diagram and re-equilibrate in to a dense material. **C)** Experimental data for the lens system plotted in context of patchy colloid theory. Likely bond temperatures for the system are shown in the height of the colored squares, while the density spanned by each experimentally characterized layer is shown in the width. Spinodal lines from theory (11) are shown in shades of grey. Circles underneath the x-axis indicate the densities of the dense material generated dilution, the radius from which samples were taken is shown inside.

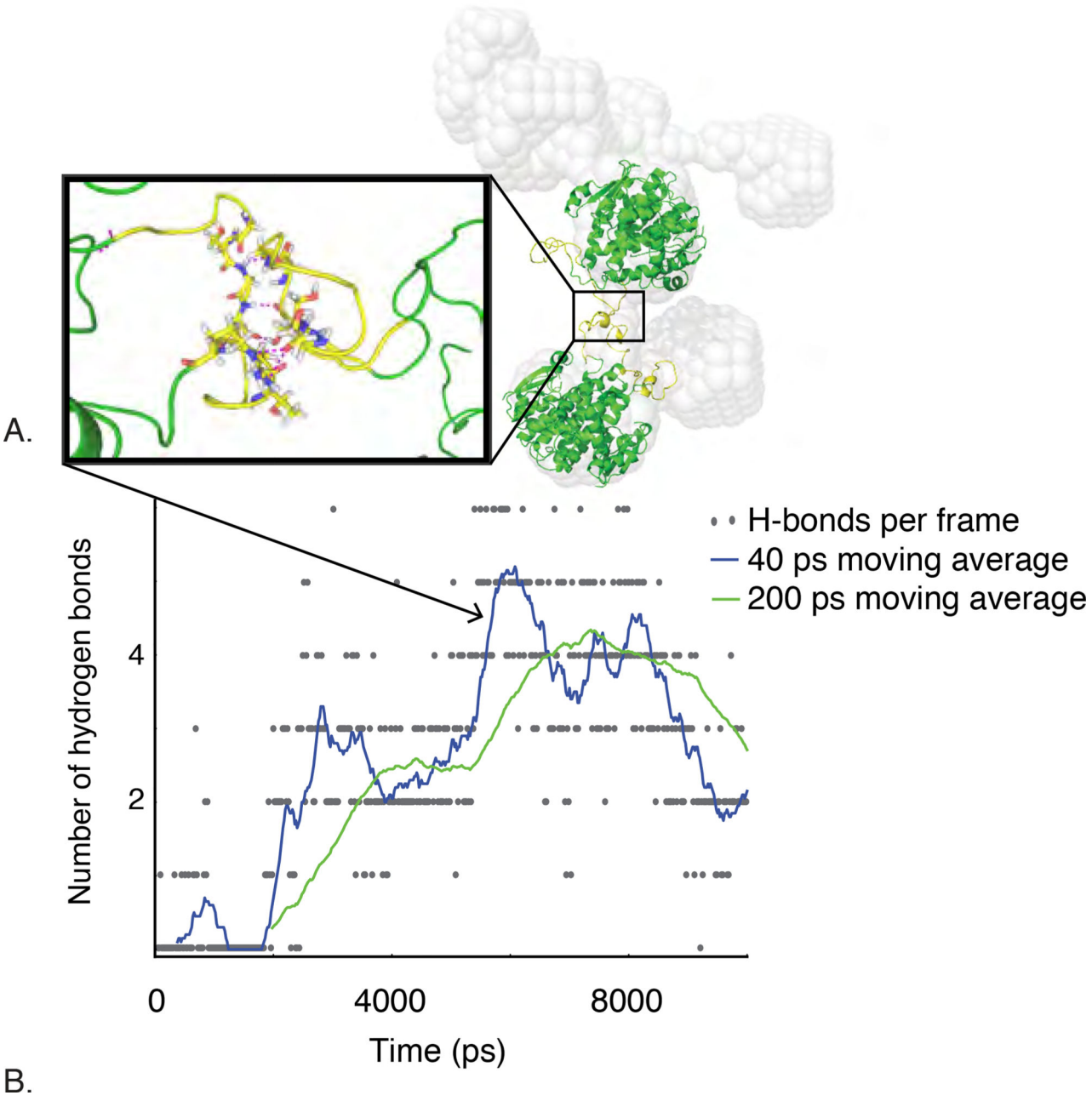

Figure 5.

MD simulation of patch-like loop-loop interaction in low-index regions of the lens. **A)** Homology models of S-crystallin dimers, shown in green and yellow, with centers of mass positioned according to DAMMIF output, shown in grey. Inset shows the interaction of the two unstructured loops of proteins in this orientation after ~6,000 ps of MD simulation. Hydrogen bonds between the two loops are shown in pink. **B)** Number of hydrogen bonds between two loops binding together two separate S-crystallin dimers in MD simulation as a

function of time. Hydrogen bonds form between the two loops after a few hundred ps and increase to a maximum of six, with a 200-ps moving average of four.