Increased hydrophobic surface exposure in the cataract-related G18V variant of human γ S-crystallin

Domarin Khago,^{†,‡} Eric K. Wong,^{†,‡} Carolyn N. Kingsley,[†] J. Alfredo Freites,[†]
Douglas J. Tobias,^{*,†} and Rachel W. Martin^{*,†,¶,‡}

*Department of Chemistry, University of California, Irvine, Irvine, CA 92697, , and
Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine
CA 92697*

E-mail: dtobias@uci.edu; rwmartin@uci.edu

Abstract

Background: The objective of this study was to determine whether the cataract-related G18V variant of human γ S-crystallin has increased exposure of hydrophobic residues that could explain its aggregation propensity and/or recognition by α B-crystallin.

Methods: We used an ANS fluorescence assay and NMR chemical shift perturbation to experimentally probe exposed hydrophobic surfaces. These results were compared to flexible docking simulations of ANS molecules to the proteins, starting with the solution-state NMR structures of γ S-WT and γ S-G18V.

*To whom correspondence should be addressed

[†]Department of Chemistry, University of California, Irvine, Irvine, CA 92697

[‡]These authors contributed equal work.

[¶]Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine CA 92697

Results: γ S-G18V exhibits increased ANS fluorescence, suggesting increased exposed hydrophobic surface area. The specific residues involved in ANS binding were mapped by NMR chemical shift perturbation assays, revealing ANS binding sites in γ S-G18V that are not present in γ S-WT. Molecular docking predicts three binding sites that are specific to γ S-G18V corresponding to the exposure of a hydrophobic cavity located at the interdomain interface, as well as two hydrophobic patches near a disordered loop containing solvent-exposed cysteines, all but one of which is buried in γ S-WT.

Conclusions: Although both proteins display non-specific binding, more residues are involved in ANS binding to γ S-G18V, and the affected residues are localized in the N-terminal domain and the nearby interdomain interface, proximal to the mutation site.

General Significance: Characterization of changes in exposed hydrophobic surface area between wild-type and variant proteins can help elucidate the mechanisms of aggregation propensity and chaperone recognition, presented here in the context of cataract formation. Experimental data and simulations provide complementary views of the interactions between proteins and the small molecule probes commonly used to study aggregation.

Highlights: Chemical shift perturbations reveal ANS bindings local to the N-terminal domain and the adjacent interdomain interface.

Residue contacts from rigid docking are consistent with the solution NMR data.

γ S-G18V exhibits unique exposure of hydrophobic patches located in a disordered loop region containing exposed cysteines as well as a large hydrophobic cavity located at the interdomain interface in the variant.

Keywords: Structural crystallin; cataract; protein aggregation; chemical shift perturbation; ANS binding assay; docking

Introduction

High concentrations of closely packed crystallin proteins are necessary for maintaining the transparency and refractive index gradient of the eye lens. The human lens has several structural crystallins that are found with different radial distributions; the focus of this study is γ S-crystallin, which is preferentially located in the lens cortex (periphery).^{1,2} The solution-state NMR structure of wild-type γ S-crystallin has been determined,³ revealing a double Greek key architecture for each of the two domains, consistent with the structures of other $\beta\gamma$ -crystallins. The childhood-onset cataract variant G18V (γ S-G18V) is structurally similar to γ S-WT, but it has dramatically lower thermal stability and solubility,^{4,5} as well as strong, specific interactions with α B-crystallin, the holdase chaperone of the lens.³ Despite the well-documented aggregation propensity and reduced stability of γ S-G18V, the particular intermolecular interactions leading to its aggregation are as yet unknown. Protein self-aggregation leading to cataract can occur due to an increase in net hydrophobic interactions, as previously shown in the congenital Coppock-type cataract variant D26G γ S-crystallin,⁶ the cerulean cataract variant P23T γ D-crystallin,⁷ acetylation of the G1 and K2 residues in γ D-crystallin,⁸ and the lamellar cataract variant D140N α B-crystallin.⁹ All of these mutations introduce altered conformations that produce lowered solubility by exposure of hydrophobic patches on the surface, even though the structural differences from their wild-type counterparts are relatively subtle. γ S-G18V is no exception; the mutation does not cause large-scale unfolding or rearrangement into a misfolded conformation, but rather produces altered intermolecular interactions with itself and with α B-crystallin.³

The fluorescent probe 1-anilinonaphthalene-8-sulfonate (ANS), which has both negatively charged and hydrophobic moieties, is often used to quantify exposed hydrophobic surface patches in proteins by introducing known concentrations of ANS into a protein solution and measuring its emission spectrum.^{7,10,11} Two types of protein-ANS interactions are required for fluorescence enhancement: hydrophobic interactions between the conjugated

ring system of ANS and the protein surface,¹² and electrostatic interactions between the sulfonate group and positively charged side chains at the binding site.¹³ An increase in fluorescence intensity indicates that either more ANS is binding to the protein surface, or that it is bound more tightly, correlating with higher surface hydrophobicity. This method has been used to characterize exposed hydrophobic surface in a number of protein systems, including the mitochondrial chaperone protein Atp11p, which recognizes its client proteins via hydrophobic interactions,¹⁴ and aggregation-prone variants of superoxide dismutase-1 (SOD1), an essential cellular enzyme whose aggregation is associated with amyotrophic lateral sclerosis (ALS).^{15,16} Despite the utility of ANS binding as a probe of hydrophobic surface exposure, and the sensitivity afforded by using fluorescence as a reporter, this assay is limited by the lack of detailed information about which amino acid residues, or even general regions of the protein, are taking part in the dye-binding interaction. NMR chemical shift perturbation (CSP) mapping can forge a link between fluorescence enhancement upon dye binding and the corresponding changes in the local chemical environment of specific residues in the protein. Comparisons between wild-type and variant proteins can then be used to compare differences in exposure of hydrophobic residues on the surface under particular solution conditions. CSP mapping is a commonly used technique for investigating protein-protein or protein-ligand interactions and interfaces,¹⁷ and is the basis of the “SAR by NMR” methodology that is indispensable in the identification of active pharmaceutical agents.¹⁸

Molecular docking, a computational technique used widely to model the conformation of protein-ligand complexes, enables experimental perturbations to be analyzed in atomistic detail. Bound ligand conformations, or poses, are ranked using an empirical scoring function designed to evaluate intermolecular interactions using minimal computational time. Conventionally, knowledge of the active site is used to guide the pose generation, often in the context of screening large libraries of compounds against known protein structures.^{19–22} However, docking protocols without prior knowledge of the active site (blind docking),²³

have successfully identified putative allosteric binding sites of drugs, leading to the design of novel allosteric modulators,²⁴ and fluorescent dyes.^{10,25} Bis-ANS binding sites found by docking, validated with steady-state and time-resolved fluorescence assays, have been used to identify hydrophobic patches in a lipase from *Bacillus subtilis*.²⁶

Materials and Methods

ANS fluorescence assay

Wild-type and G18V γ S-crystallins were expressed and purified as described previously.⁵ Fluorescence spectra were collected as a function of ANS concentration for γ S-WT and γ S-G18V with a F4500 Hitachi fluorescence spectrophotometer. The excitation and emission wavelengths were 390 nm and 500 nm, respectively, with slits set to 5 nm. Protein concentrations for both γ S-WT and γ S-G18V were approximately 1 mg/mL in 10 mM sodium phosphate buffer and 0.05% sodium azide at pH 6.9. ANS concentrations ranging from 5 μ M to 2 mM were measured using $\epsilon = 4.95 \text{ mM}^{-1} \text{ cm}^{-1}$ at 350 nm.²⁷

NMR sample preparation

Purified protein with the 6x-His tag removed was concentrated and supplemented with 2 mM TMSP, 10% D₂O, and 0.05% sodium azide. The final concentration of all γ S-WT and γ S-G18V samples was 0.3 mM. ANS was titrated into the protein samples to give final molar ratios of 1:0, 1:0.5, 1:1, and 1:2 of γ S:ANS. Spectra were acquired at 25 °C.

NMR experiments

Experiments were performed on a Varian ^{Unity}INOVA spectrometer (Agilent Technologies) operating at 800 MHz and equipped with a ¹H–¹³C–¹⁵N 5 mm tri-axis PFG triple-resonance probe, using an 18.8 Tesla superconducting electromagnet (Oxford instruments). Decoupling

of ^{15}N nuclei was performed using the GARP sequence.²⁸ ^1H chemical shifts were referenced to TMSP, and ^{15}N shifts were referenced indirectly to TMSP. NMR data were processed using NMRPipe²⁹ and analyzed using CcpNMR Analysis.³⁰ Center operating frequencies and (unless otherwise stated) center frequency offsets were as follows:

Center	^1H : 799.8056964 MHz	^{13}C : 201.1282461 MHz	^{15}N : 81.0504078 MHz
Offset	^1H : -294.932 Hz (4.8 ppm)	^{13}C : -9863.17 Hz (43 ppm)	^{15}N : 2400 Hz (116.7 ppm)

Calculation of chemical shift perturbations

^1H - ^{15}N HSQC spectra of γS -WT and γS -G18V were collected in the presence and absence of ANS at concentration ratios of 1:0, 1:0.5, 1:1, and 1:2 of γS :ANS, and resonances were identified and assigned based on chemical shift data previously collected by our group. Resonances showed perturbations that are indicative of ANS binding. The change in chemical shift for each peak in the 2D spectrum upon ANS binding was calculated using the following chemical shift perturbation (CSP) equation:

$$\Delta\delta_{avg} = \sqrt{\frac{(\Delta\delta_N/5)^2 + (\Delta\delta_H)^2}{2}} \quad (1)$$

A strong-binding threshold for each set of conditions was set at two times the root mean square (RMS) of the calculated CSP, while the weak-binding threshold was set at half the RMS to determine which residues had strong or weak binding with ANS. The values used for each threshold appear in Supplementary Table S1.

Binding site search by rigid receptor docking

Protein coordinates were obtained from the NMR structures of γS -WT and γS -G18V crystallins (PDB ID: 2M3T and 2M3U).³ Autodock Tools³¹ was used to prepare both the receptor (crystallin) and ligand (ANS) by merging non-polar hydrogens atoms into united heavy atoms. Gasteiger charges³² were added to each atom. The sulfonic acid group of ANS was deprotonated before processing by Autodock Tools. Molecular docking was performed using

Autodock Vina.³³ In order to ensure good coverage of the protein binding surface, 27 search spaces were placed in an overlapping 3 x 3 x 3 grid around the protein (Supplementary Figure S1). Since Autodock Vina works optimally with search spaces with at most a 27,000 Å³ volume, a 30 x 30 x 30 Å search space was chosen. The exhaustiveness parameter was set to 20 (over the default value of 8) in order to ensure an extensive search of the protein surface. Docking was performed over each one of twenty solution-state NMR conformations for either γ S-WT or γ S-G18V. The resulting poses were screened to ensure that both electrostatic and hydrophobic interactions required for ANS fluorescence enhancement upon binding were present. Docked poses that did not include both interactions within the first coordination shell of the ANS-protein radial distribution function were considered non-fluorescent and removed from the docked set. The screened docked set covers most of the protein surface (see Supplementary Information Figure S2A).

Calculation of residue contacts

To compare the screened docked set with the residue-based CSP data, ANS-residue contact frequencies were calculated by summing the Boltzmann weights of all the poses in contact with a given residue. The Boltzmann weight of a given docked pose was calculated according to

$$w_i = \frac{\exp(-E_i/k_B T)}{\sum_i \exp(-E_i/k_B T)} \quad (2)$$

where i is the index of the docked pose, E_i is the pose binding energy, k_B is the Boltzmann constant, and T is the absolute temperature. The residue contact frequencies for each protein are shown in Supplementary Figure S3. Following the CSP analysis, to determine which residues had strong or weak binding with ANS, a strong-binding threshold was set at two times the RMS of the calculated ANS-residue contact frequency, while the weak-binding threshold was set at the RMS value. The values used for each threshold appear in the Supplementary Information (Table S1).

Flexible refinement of binding sites

A flexible docking refinement was performed near all the highly perturbed residues identified using the strong-binding cutoff on the CSP data described in the previous section. Docking search spaces were defined by clustered conformations of ANS from the screened docked set used to calculate the ANS-residue contact frequencies. Using a root-mean-square deviation cutoff (RMSD) of 5.0 Å, clustered poses were grouped into potential binding sites near the experimentally perturbed residues (see Supplementary Figure S2B). Search spaces were defined as boxes surrounding the clustered ligands with an 8 Å padding. The padding was necessary to include flexible side chains within the search space. Residues with an experimental CSP above the weak-binding cutoff were considered as flexible. A total of five potential binding sites were used to dock ANS to either flexible γ S-WT or γ S-G18V. The resulting poses were clustered again, and the locations and interactions of each pose were compared visually.

Results and Discussion

ANS fluorescence indicates that the relative surface hydrophobicity of γ S-G18V is higher than that of γ S-WT

Dye-binding assays were performed on γ S-WT and the aggregation prone variant, γ S-G18V. The ANS fluorescence measurements for γ S-WT and γ S-G18V, shown in Figure 1, indicate more exposed hydrophobic surface in γ S-G18V compared to its wild type counterpart. These data also allow determination of the lowest ANS concentration required to produce the maximum emission before saturation, which was 1.5 mM for γ S-WT and 1 mM for γ S-G18V. The lower concentration required to saturate γ S-G18V is consistent with the observation that it binds ANS more readily than wild-type.

Chemical shift perturbation mapping reveals the residues involved in ANS binding and the relative strengths of the interactions

Binding interactions between ANS and γ S-WT or γ S-G18V were measured at concentration ratios of 1:0, 1:0.5, 1:1, and 1:2 of γ S:ANS, using CSP mapping via ^1H - ^{15}N HSQC spectra. Selected regions of the NMR spectra where resonances show perturbations indicative of ANS binding are shown in Figure 2. The full NMR spectra can be found in the Supplemental Information (Supplementary Figures S5 and S6). The change in chemical shift for each peak in the 2D spectrum upon ANS binding was calculated using Equation 1. Representative CSP data for γ S-WT and γ S-G18V upon 1:1 ANS binding are shown in Figure 3. The complete set of calculated CSP data can be found in the Supplemental Information (Supplementary Figures S7 and S8). While nonspecific binding is observed throughout the surfaces of both proteins, γ S-G18V binds ANS more strongly in the N-terminal domain (approximately the first 100 residues). The maximum ANS binding occurs within residues 15 through 50, close to the mutation site. These observations are mapped onto the protein structures in Figure 4 (left panel) where the residues exhibiting strong (CSP at least two times the RMS) and weak (CSP at least half the RMS) ANS binding are highlighted. For γ S-WT, strong binding residues are highlighted in bright green and weak binding residues in pale green. For γ S-G18V, strong binding residues are highlighted in dark blue and weak binding residues in pale blue. Although some strong binding residues are observed in both proteins near the mutation site, (e.g. G18 in γ S-WT and D22 in γ S-G18V), G18V displays more ANS binding, both strong and weak, in the N-terminal domain. Strong binding is also seen in the interdomain interface of γ S-WT, (residues L62, S82, and H123), and γ S-G18V (residues L62, W73, H87, L88, and G91). These results are consistent with the observation that α B-crystallin strongly binds near the N-terminal domain and the interdomain interface in γ S-G18V, but not γ S-WT.³ Thus, the ANS-binding data support the hypothesis that the chaperone may be recognizing an exposed hydrophobic patch in this region of γ S-G18V.

In order to characterize the aggregation states of γ S-WT and γ S-G18V, dynamic light scattering (DLS) data were acquired for both proteins under the same solution conditions used in the NMR experiments (data shown in the Supplementary Information Figure S4). As observed in previous studies of γ S-G18V,^{3,5} the γ S-WT solution contains only monomers, while γ S-G18V shows a slightly broader range of sizes consistent with transient formation of dimers and potentially other small oligomers. However, the NMR spectra rule out the presence of a significant stable population of large aggregates; the linewidths for representative peaks in the HSQC spectra are comparable to the corresponding peaks in γ S-WT. Linewidth comparisons for representative peaks in the spectra of γ S-WT and γ S-G18V are tabulated in the Supplemental Information (Table S2). Despite the presence of transient oligomers in the γ S-G18V sample, consistent with its increased aggregation propensity, it is clear from the linewidth data that the chemical shift changes upon addition of ANS are due to dye binding rather than a change in aggregation state. Although the oligomerization states of the starting solutions were not identical, this is accounted for by the chemical shift differences between γ S-WT and γ S-G18V in the absence of ANS, while the chemical shift perturbations reflect binding of each protein to ANS. If stable, large complexes were present in the NMR samples, the increased aggregation would be expected to cause significant line-broadening and disappearance of signals, as was observed for mixtures of γ S-G18V with α B-crystallin,³ where large complexes were formed and TROSY techniques were required to observe the NMR signals. Although it is possible to prepare purely monomeric samples of γ S-G18V at low pH, for the current study, neutral pH was chosen in order to investigate intermolecular interactions under more physiologically realistic conditions.

In order to investigate whether ANS changes the oligomerization states of γ S- α B complexes and interferes with binding of α B-crystallin to γ S-G18V, we performed gel filtration chromatography (Figure 5). Samples of γ S-WT and γ S-G18V were prepared at 1 mg/mL and compared to equivalent samples in the presence of α B-crystallin (1:1) and both α B-crystallin

and ANS (1:1:1). For γ S-WT alone, the sample is mostly monomeric with a small amount of dimers. Upon addition of α B-crystallin, a population of larger oligomers at about 160 kDa appears, at the expense of the populations of both the monomeric and dimeric states. Addition of ANS to this mixture slightly increases the proportion of large aggregates. In the case of γ S-G18V, both the initial oligomerization states and the effect of adding ANS is different. Initially, although much of the sample is monomeric, small populations of dimers and large oligomers exist. The peak at 160 kDa is much broader than in γ S-WT, suggesting greater polydispersity. In the presence of α B-crystallin, the main effect is a dramatic narrowing of the peak corresponding to large oligomers, indicating a more uniform population. Addition of ANS to this mixture produces both further narrowing and an increase in the population of monomers, suggesting that interaction with ANS does disrupt the α B- γ S complex formation. The full chromatogram including molecular weight standards is provided in the Supplemental Information (Figure S9).

Docking of ANS on the protein surface predicts more binding sites on γ S-G18V than γ S-WT and allows interpretation of the CSP data

Rigid receptor docking resulted in a total of 4860 docked poses (27 search spaces $\times$ 20 NMR conformations $\times$ 9 poses/search space). After screening for poses consistent with ANS fluorescence enhancement upon binding, 3423 poses and 3367 poses remained for γ S-WT and γ S-G18V, respectively (see Supplementary Information Figure S2A). Filtered poses covered nearly the entire surface of the protein and exhibit a broad range of scores (from -2 kcal/mol to -7 kcal/mol, with a mean of -4.5 kcal/mol). Due to the pocket-like shape of the interdomain interface, ANS preferentially bound to the large hydrophobic pocket between the N- and C-terminal domains. However, sites were identified near all highly perturbed residues with comparable binding scores (see Supplementary Information Figure S3). Flexible docking poses located near the highly perturbed residues according to the CSP data had binding

scores between -4.5 kcal/mol and -6.0 kcal/mol, consistent with a stronger preference for ANS to bind near the perturbed residues. A total of ten binding sites were found for γ S-G18V and nine binding sites for γ S-WT using flexible docking. Most of these binding sites were very similar in both γ S-WT and γ S-G18V. However, three binding modes were unique to γ S-G18V. The first and most populated binding mode is located in the hydrophobic cavity at the interface between the N- and C-terminal domains, shown in Figure 6A and 6D.^{34,35} Although this binding site was found in both γ S-WT and γ S-G18V, the presence of the R84-D153 salt-bridge blocks the exposure of the hydrophobic surface in γ S-WT. In contrast, γ S-G18V lacks this salt-bridge interaction, which exposes the interdomain hydrophobic cavity and allows the entry of ANS into the interdomain binding site. This is consistent with the experimental NMR data, which indicate that chemically perturbed residues, H87 and L88, located near the interdomain pose, interact strongly with ANS only in γ S-G18V (Figure 6). The second and third binding sites are located close to residues 20 - 30, which includes a loop region containing three cysteine residues (C23, C25, and C27). As a result of the G18V mutation, C23 and C27 become solvent exposed, suggesting possible formation of intermolecular disulfide bridges, consistent with the observation that an excess of reducing agents abrogates the formation of small oligomers.³⁶ Previous studies suggested that the exposure of these cysteines results from a disruption in secondary structure due to the burial of V18 side chain.³ As a result of this cysteine exposure and concomitant structural changes, a new hydrophobic pocket is uncovered as the second ANS binding site. Although ANS binds this Cys loop in γ S-WT after flexible docking refinement, it is not in direct contact with any hydrophobic surface, suggesting that the pose may not be consistent with and enhancement in ANS fluorescence (Figure 6B). In contrast, when the hydrophobic pocket is exposed, as it is in γ S-G18V, ANS becomes buried deep within the pocket (Figure 6E). This conformation provides both the hydrophobic interactions necessary for fluorescence as well as reduced quenching due to water exposure.¹² In addition to cysteine exposure, the third binding site shows additional hydrophobic surface exposure due to the cysteine loop

separating from the main Greek key motif. This binding site was not found in γ S-WT using the same docking search space, indicating that this hydrophobic patch is a unique characteristic of γ S-G18V (Figure 6F). Additionally, the CSP data shows local perturbation of the backbone amides of the residues involved in these three binding sites only for γ S-G18V. The presence of these γ S-G18V-specific binding sites can explain the higher ANS fluorescence intensity of the variant protein over WT, and they also identify exposed hydrophobic patches which may potentially serve as protein-protein interfaces in crystallin aggregates, and which can be targeted in future mutagenesis studies.

The CSP data and the ANS-residue contact data from the docking simulations show generally good agreement in that the same protein regions were observed to bind ANS (see Figure 4). In some cases, the specific residues classified as strong binding vary between experimental and docking results, but coverage of both strong- and weak-binding residues are nearly identical (highlighted in dark green for γ S-WT, and dark blue γ S-G18V in the right panel of Figure 4). This outcome is to be expected because the docking scoring function is more effective at identifying binding sites than distinguishing more subtle changes in binding energy: the standard error of the Autodock Vina scoring function³³ is larger than the variation among scored poses. The agreement between rigid protein docking results and experimental ANS binding results suggests that there is no major change in protein conformation upon binding of ANS, supporting the hypothesis that hydrophobic patches on the surface are involved in intermolecular interactions. Good agreement between the experimental and docking results further confirms that ANS binding is localized near the mutation site in the N-terminal domain for γ S-G18V, consistent with the CSP data. Experimental and simulation results are also consistent on the binding of ANS to the exposed interdomain hydrophobic surface located in the interdomain interface between the two domains due to the breaking of the R84-D153 salt bridge in γ S-G18V. Exposure of this hydrophobic patch facilitates ANS binding and may be involved in hydrophobic protein-protein interactions.

Conclusion

In this study, we have used ANS fluorescence, solution-state NMR chemical shift perturbation mapping, and molecular docking to investigate the differences in exposed hydrophobic surface between human γ S-crystallin and its cataract-related γ S-G18V variant. The experimental results indicated that both proteins have a fairly high level of nonspecific binding, but both the fluorescence and NMR measurements indicate more ANS binding to γ S-G18V, particularly in the N-terminal domain near the mutation site. The docking studies, in agreement with the NMR data, found three binding modes that are unique to γ S-G18V that were not found in γ S-WT: one in the exposed hydrophobic patch in the interdomain interface and two binding modes in the exposed hydrophobic pocket formed when the cysteine loop becomes solvent exposed. Using docking and binding assays, hydrophobic surface patches were identified that may be responsible for some of the intermolecular interactions between crystallins that promote aggregation in the lens. The results may also guide the design of future mutagenesis and drug-binding studies to further investigate the importance of such intermolecular interactions in mediating γ S-crystallin solubility and aggregation resistance in the healthy eye lens.

Acknowledgement

The authors thank Dr. Philip Dennison and Dr. Dmitry Fishman for excellent management of the UCI Chemistry Department’s NMR and Laser Spectroscopy Facilities, and Anne Diehl and Harmut Oschkinat for providing the α B-crystallin sample. The docking calculations were performed on the Extreme Science and Engineering Discovery Environment (XSEDE),³⁷ which is supported by National Science Foundation grant ACI-1053575. The computational work was supported by National Science Foundation grant DMR-1410415 to RWM and DJT, and the experimental work was supported by National Institutes of Health R01EY021514 to

RWM.

Figure 1: A. Molecular surface representation of γ S-WT (green) and γ S-G18V (blue) based on the solution-state NMR structures (PDB ID 2M3T and 2M3U, respectively). Hydrophobic residues are highlighted in orange. B. Fluorescence spectra representing ANS binding monitored at 500 nm using γ S-WT and γ S-G18V crystallins. Protein concentrations for both γ S-WT and γ S-G18V were approximately 1 mg/mL. Saturation occurred at 1.5 mM ANS for γ S-WT and 1 mM ANS for γ S-G18V. Higher emission was observed for γ S-G18V, indicating more hydrophobic surface area exposed than for γ S-WT.

Figure 2: Selected portions of the overlaid ^1H - ^{15}N HSQC spectra of γ S-WT and γ S-G18V. Experiments were carried out using ratios of 1:0, 1:0.5, 1:1, and 1:2 of γ S:ANS. Spectra were acquired at 25 °C with concentrations of all γ S-WT and γ S-G18V samples at 0.3 mM. Resonances having a change in chemical shift indicate ANS binding to specific residues, which is quantified using the chemical shift perturbation equation (Equation 1).

Figure 3: Average chemical shift perturbation (CSP) of γ S-WT (green) and γ S-G18V (blue). Nonspecific binding, with maximum perturbation in the N-terminal domain, is observed in both proteins. However, in γ S-G18V more of the CSPs are localized to the N-terminal domain, particularly between residues 15 to 50, in the cysteine loop near the mutation site. Inspection of the structures confirms that this region is exposed to solvent in γ S-G18V but not γ S-WT.

Figure 4: ANS interactions with γ S-WT and γ S-G18V. The strong-binding threshold and weak-binding threshold were defined as two times the RMS and half the RMS, respectively. Experimental CSP values indicate that ANS binding occurs throughout the N- and C-terminal domains for γ S-WT, (strong binding residues in green and weak binding residues in pale green), while in γ S-G18V ANS binding mainly occurs at the N-terminal domain (strong binding residues in blue and weak binding residues in pale blue). Some strong binding is observed in the N-terminal domain for both proteins near the mutation site, e.g. G18 in γ S-WT and D22 in γ S-G18V. However, G18V displays more ANS binding (both strong and weak) overall in the N-terminal domain. Strong binding is also observed in the interdomain interface of γ S-WT, residues L62, S82, and H123, and γ S-G18V, residues L62, W73, H87, L88, and G91. G18V exhibits more binding (strong and weak) within that interdomain interface suggesting that this variant has higher surface hydrophobicity localized to the N-terminal domain near the mutation site and the interdomain interface. Coverage of both strong and weak binding residues are nearly identical between experimental and docking results, highlighted in dark green for γ S-WT and dark blue for γ S-G18V, indicating that the docking results are in good agreement with the experimental data.

Figure 5: Gel filtration chromatograms for γ S-WT (A) and γ S-G18V (B) in the presence and absence of α B-crystallin and ANS. (A) For γ S-WT alone (green), the sample is mostly monomeric (10 kDa) with a small amount of dimers (22 kDa). Upon addition of α B-crystallin (orange), a population of larger oligomers (160 kDa) appears, at the expense of the populations of both the monomeric and dimeric states. Addition of ANS to this mixture (red) slightly increases the proportion of large aggregates. (B) For γ S-G18V alone (blue), much of the sample is monomeric, although small populations of dimers and large oligomers exist. The peak at 160 kDa is much broader than in γ S-WT, suggesting greater polydispersity. In the presence of α B-crystallin (cyan), the main effect is a dramatic narrowing of the peak corresponding to large oligomers, indicating a more uniform population. Addition of ANS to this mixture (purple) produces both further narrowing and an increase in the population of monomers, suggesting that interaction with ANS does disrupt the α B- γ S complex formation.

Figure 6: Docking poses of ANS bound to a flexible γ S-WT and γ S-G18V receptor. Three binding sites were found to be unique to γ S-G18V. The protein surface was generated with MSMS;³⁸ red, blue, green, and white correspond to negative, positive, polar and hydrophobic regions, respectively. ANS is shown in licorice representation. The R84-D153 salt bridge and cysteine residues (C23, C25, and C27) critical to the hydrophobic patch availability are shown in space-filling representation. In the left-hand panels, residues defined as strong/weak binding by CSP data have their backbone amides represented as spheres. Large spheres represent strongly-binding residues, and small spheres represent weakly-binding residues. Atoms are colored by element (carbon, cyan; nitrogen, blue; oxygen, red; hydrogen, white; sulfur, yellow). (A & D) At site 1, the R84-D153 salt bridge separates to expose the hydrophobic cavity at the interdomain interface. Although a pose is generated in both proteins, the lack of perturbed residues at the binding site, according to the CSP data, indicates that the binding site is inaccessible to ANS in γ S-WT. (B & E) At site 2, the docked pose of ANS shifts from a polar surface in γ S-WT to inserting into a hydrophobic cavity in γ S-G18V. Due to specific backbone torsions propagating from the G18V mutation site that keep the V18 buried, C23 and C27 become solvent exposed and reveal a hydrophobic cavity. This pose is located near the largest perturbed residue in γ S-G18V according to the CSP data. (C & F) At site 3, the disordered cysteine loop separates from the Greek key motif of the N-terminal domain resulting in exposure of an additional hydrophobic patch. No equivalent pose could be generated on the γ S-WT structure.

References

- (1) Aarts, H. J.; Lubsen, N. H.; Schoenmakers, J. G. Crystallin gene expression during rat lens development. *European Journal of Biochemistry* **1989**, *183*, 31–36.
- (2) Nagineni, C. N.; Bhat, S. P. Lens fiber cell differentiation and expression of crystallins in co-cultures of human fetal lens epithelial cells and fibroblasts. *Experimental Eye Research* **1992**, *54*, 193–200.
- (3) Kingsley, C. N.; Brubaker, W. D.; Markovic, S.; Diehl, A.; Brindley, A. J.; Oschkinat, H.; Martin, R. W. Preferential, specific binding of human α B-crystallin to a cataract-related variant of γ S-crystallin. *Structure* **2013**, *21*, 2221–2227.
- (4) Ma, Z.; Piszczek, G.; Wingfield, P.; Sergeev, Y.; Hejtmancik, J. The G18V CRYGS mutation associated with human cataracts increases γ S-crystallin sensitivity to thermal and chemical stress. *Biochemistry* **2009**, *48*, 7334–7341.
- (5) Brubaker, W. D.; Freites, J. A.; Golchert, K. J.; Shapiro, R. A.; Morikis, V.; Tobias, D. J.; Martin, R. W. Separating instability from aggregation propensity in γ S-crystallin variants. *Biophys. J.* **2011**, *100*, 498–506.
- (6) Karri, S.; Kasetti, R. B.; Vendra, V. P. R.; Chandani, S.; Balasubramanian, D. Structural analysis of the mutant protein D26G of human γ S-crystallin, associated with Coppock cataract. *Molecular Vision* **2013**, *19*, 1231–1237.
- (7) Pande, A.; Ghosh, K. S.; Banjeree, P. R.; Pande, J. Increase in surface hydrophobicity of the cataract-associated P23T mutant of human γ D-crystallin is responsible for its dramatically lower retrograde solubility. *Biochemistry* **2010**, *49*, 6122–6129.
- (8) DiMauro, M. A.; Nandi, S. K.; Raghavan, C. T.; Kar, R. K.; Wang, B.; Bhunia, A.; Nagaraj, R. H.; Biswas, A. Acetylation of Gly1 and Lys2 promotes aggregation of human γ D-crystallin. *Biochemistry* **2014**, *53*, 7269–7282.

- (9) Liu, Y.; Zhang, X.; Luo, L.; Wu, M.; Zeng, R.; Cheng, G.; Hu, B.; Liu, B.; Liang, J.; Shang, F. A novel α B-crystallin mutation associated with autosomal dominant congenital lamellar cataract. *Investigative Ophthalmology & Visual Science* **2006**, *47*, 1069–1075.
- (10) Banerjee, P. R.; Puttamadappa, S. S.; Pande, A.; Shekhtman, A.; Pande, J. Increased Hydrophobicity and Decreased Backbone Flexibility Explain the Lower Solubility of a Cataract-Linked Mutant of γ D-Crystallin. *Journal of Molecular Biology* **2011**, *412*, 647–659.
- (11) Bharat, S. V.; Shekhtman, A.; Pande, J. The cataract-associated V41M mutant of human γ S-crystallin shows specific structural changes that directly enhance local surface hydrophobicity. *Biochemical and Biophysical Research Communications* **2014**, *443*, 110–114.
- (12) Matulis, D.; Lovrien, R. 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation. *Biophysical Journal* **1998**, *74*, 422–429.
- (13) Ory, J. J.; Banaszak, L. J. Studies of the ligand binding reaction of adipocyte lipid binding protein using the fluorescent probe 1, 8-anilinonaphthalene-8-sulfonate. *Biophysical Journal* **1999**, *77*, 1107–1116.
- (14) Sheluho, D.; Ackerman, S. H. An accessible hydrophobic surface is a key element of the molecular chaperone action of Atp11p. *Journal of Biological Chemistry* **2001**, *43*, 39945–39949.
- (15) Münch, C.; Bertolotti, A. Exposure of hydrophobic surfaces initiates aggregation of diverse ALS-causing superoxide dismutase-1 mutants. *Journal of Molecular Biology* **2010**, *399*, 512–525.
- (16) Tiwari, A.; Liba, A.; Sohn, S. H.; Seetharanab, S. V.; Bilsel, O.; Matthews, C. R.; Hart, P. J.; Valentine, J. S.; Hayward, L. J. Metal deficiency increases aberrant hy-

- drophobicity of mutant superoxide dismutases that cause amyotrophic lateral sclerosis. *Journal of Biological Chemistry* **2009**, *284*.
- (17) Williamson, M. P. Using chemical shift perturbation to characterize ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2013**, *73*, 1–16.
- (18) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. Discovering high affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.
- (19) ten Brink, T.; Aguirre, C.; Exner, T. E.; Krimm, I. Performance of protein-ligand docking with simulated chemical shift perturbations. *Journal of Chemical Information and Modeling* **2015**, *55*, 275–283.
- (20) Huang, S.; Zou, X. Efficient molecular docking of NMR structures: Application to HIV-1 protease. *Protein Science* **2007**, *16*, 43–51.
- (21) Stark, J. L.; Powers, R. Application of NMR and molecular docking in structure-based drug discovery. *Topics in Current Chemistry* **2011**, *326*, 1–34.
- (22) van Dijk, A. J.; Boelens, R.; Bonvin, A. M. Data-driven docking for the study of biomolecular complexes. *FEBS Journal* **2005**, *272*, 293–312.
- (23) Hetényi, C.; van der Spoel, D. Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein science : a publication of the Protein Society* **2002**, *11*, 1729–1737.
- (24) Iorga, B.; Herlem, D.; Barré, E.; Guillou, C. Acetylcholine nicotinic receptors: Finding the putative binding site of allosteric modulators using the “blind docking” approach. *Journal of Molecular Modeling* **2006**, *12*, 366–372.
- (25) Konuma, T.; Lee, Y.-H.; Goto, Y.; Sakurai, K. Principal component analysis of chemical shift perturbation data of a multiple-ligand-binding system for elucidation of respective binding mechanism. *Proteins* **2013**, *81*, 107–118.

- (26) Kamal, M. Z.; Ali, J.; Rao, N. M. Binding of bis-ANS to Bacillus subtilis lipase: a combined computational and experimental investigation. *Biochimica et biophysica acta* **2013**, *1834*, 1501–1509.
- (27) Weber, G.; Young, L. B. Fragmentation of bovine serum albumin by pepsin. *The Journal of Biological Chemistry* **1964**, *239*, 1415–1423.
- (28) Shaka, A. J.; Barker, P. B.; Freeman, R. J. Computer-optimized decoupling scheme for wideband applications and low-level operation. *Journal of Magnetic Resonance* **1985**, *64*.
- (29) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* **1995**, *6*, 277–293.
- (30) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pjion, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* **2005**, *59*, 687–696.
- (31) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.;Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry* **2009**, *30*, 2785–2791.
- (32) Gasteiger, J.; Marsili, M. A New Model for Calculating Atomic Charges in Molecules. *Tetrahedron Lett.* **1978**, *34*, 3181–3184.
- (33) Trott, O.; Olson, A. J. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2010**,
- (34) Mills, I. A.; Flaugh, S. L.; Kosinski-Collins, M. S.; King, J. A. Folding and stability

- of the isolated Greek key domains of the long-lived human lens proteins γ D-crystallin and γ S-crystallin. *Protein Science* **2007**, *16*, 2427–2444.
- (35) Flaugh, S. L.; Kosinski-Collins, M. S.; King, J. Interdomain side-chain interactions in human γ D-crystallin influencing folding and stability. *Protein Science* **2005**, *14*, 2030–2043.
- (36) Mohr, B. G.; Dobson, C. M.; Garman, S. C.; Muthukumar, M. Electrostatic origin of in vitro aggregation of human gamma-crystallin. *Journal of Chemical Physics* **2013**, *139*, 121914.
- (37) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazelwood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; Roskies, R.; Scott, J. R.; Wilkens-Diehr, N. XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering* **2014**, *16*, 62–74.
- (38) Sanner, M. F.; Olsen, A.; Spehner, J. Fast and robust computation of molecular surfaces. Proceedings of the 11th ACM symposium on computational geometry. 1995; pp C6–C7.

Figure 1

γ S-WT

γ S-G18V

B

Figure 5

Legend

Figure 6

γ S-WT

γ S-G18V

γ S-WT

G18

γ S-G18V

V18