THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 36, pp. 19172–19183, September 2, 2016

An Internal Disulfide Locks a Misfolded Aggregation-prone Intermediate in Cataract-linked Mutants of Human γD-Crystallin*

Received for publication, May 2, 2016, and in revised form, July 12, 2016 Published, JBC Papers in Press, July 14, 2016, DOI 10.1074/jbc.M116.735977

Eugene Serebryany[‡1], Jaie C. Woodard[§1], Bharat V. Adkar[§], Mohammed Shabab[‡], Jonathan A. King[‡2], and Eugene I. Shakhnovich[§3]

From the [‡]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and the [§]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Considerable mechanistic insight has been gained into amyloid aggregation; however, a large number of non-amyloid protein aggregates are considered "amorphous," and in most cases, little is known about their mechanisms. Amorphous aggregation of γ-crystallins in the eye lens causes cataract, a widespread disease of aging. We combined simulations and experiments to study the mechanism of aggregation of two γD-crystallin mutants, W42R and W42Q: the former a congenital cataract mutation, and the latter a mimic of age-related oxidative damage. We found that formation of an internal disulfide was necessary and sufficient for aggregation under physiological conditions. Two-chain all-atom simulations predicted that one non-native disulfide in particular, between Cys^{32} and Cys^{41}, was likely to stabilize an unfolding intermediate prone to intermolecular interactions. Mass spectrometry and mutagenesis experiments confirmed the presence of this bond in the aggregates and its necessity for oxidative aggregation under physiological conditions *in vitro*. Mining the simulation data linked formation of this disulfide to extrusion of the N-terminal β-hairpin and rearrangement of the native β-sheet topology. Specific binding between the extruded hairpin and a distal β-sheet, in an intermolecular chain reaction similar to domain swapping, is the most probable mechanism of aggregate propagation.

Partially unfolded or misfolded, aggregation-prone protein conformational states are linked to a wide array of age-related protein misfolding diseases. The best studied of these conditions include amyotrophic lateral sclerosis (superoxide dismutase), Parkinson disease (α-synuclein), serpinopathies (α_1-antitrypsin), cancers with P53 and P21 tumor suppressor defects, and lens cataract (crystallins) (1–5). Some of these aggregates contain the well known amyloid structure, and others do not; even in cases where amyloid is the final aggregated state, oligomers, prefibrillar species, and amorphous aggregates are often closely linked to pathology (6, 7). Structures and interactions of specific locally unfolded or misfolded intermediate states are critically important in the mechanisms of non-amyloid aggregation (1). Here, we investigate such an intermediate-based aggregation mechanism for a lens γ-crystallin.

Because the critical properties of protein structure (thermodynamic and kinetic stability, as well as alternative, misfolded, or partially unfolded conformational states) are encoded ultimately in the sequence, point mutations have been found to underlie familial forms of virtually all misfolding diseases. However, most conformational diseases are sporadic in origin, with neither inherited nor *de novo* mutations present in the relevant genes. In the absence of mutations, such cases may arise from direct chemical modifications of residues within the relevant polypeptide. Well studied examples include superoxide dismutase 1, the eye lens crystallins, and α-synuclein (4, 8–10). In a great number of other cases, post-translational modifications cause conformational change as part of the normal function of a protein. During the course of aging, both somatic mutations and side-chain modifications tend to accumulate, further increasing the heterogeneity of the proteome (11, 12).

The human lens grows slowly in layers throughout life; its core, however, is metabolically inert and therefore provides a striking case study in protein aging (13). Lens core cells are enucleated, contain no organelles or protein synthesis and diminished degradation machinery, and are never replaced (13–15). Crystallins comprise over 90% of their total protein content. Lens transparency depends on the crystallins' structural integrity and lack of long-range packing even at very high concentrations (16). Passive chaperones, the α-crystallins, are present in lens, but their capacity declines with age, even as their substrates, the βγ-crystallins, accumulate destabilizing chemical modifications (13) due to a variety of environmental damage (17). The result is generation of partially unfolded intermediate conformational states and progressive increase in light scattering (lens turbidity) due to aggregation (18–21). No long-range structure, amyloid or otherwise, has been found in the cataractous aggregates (22, 23), except in certain rare congenital cases (24, 25), but disulfide bonds and other covalent modifications are common (13, 26–29).

Non-native disulfides can act as kinetic traps in protein folding and misfolding (30–33). Hence, some pathways of protein aggregation depend strongly on the redox environment. Nota-

* This work was supported by National Institutes of Health Grant R01 GM111955 (to E. I. S.); a National Institutes of Health Biophysics Training Grant (to J. C. W.); National Institutes of Health Grant R01 EY015834 (to J. A. K.); and fellowships from MIT Biology and the Whitehead Institute for Biomedical Research (to E. S.) The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[1] Both authors contributed equally to this work.
[2] To whom correspondence may be addressed. E-mail: jaking@mit.edu.
[3] To whom correspondence may be addressed. E-mail: shakhnovich@chemistry.harvard.edu.

ble examples include superoxide dismutase 1 (34, 35) and β_2-microglobulin (36–38). We and others have demonstrated that γ-crystallins, too, form crucial disulfide-mediated interactions. Thus, aggregation of W42Q γD-crystallin depends on internal disulfide bonds (39), and so do the folding kinetics and protease sensitivity of γB-crystallin (40). Although the crystallins are cytosolic, and therefore initially fold in a reducing environment, the redox potential of lens cytosol shifts toward increasingly oxidizing values during the course of aging and, especially, cataractogenesis (41). Accordingly, the proportion of Cys residues forming disulfide bonds increases over time (42), and mature cataracts show disulfide cross-linking in over half of the lens protein content (43).

To better understand the nature of the aggregation precursor and the mechanism of non-amyloid aggregation for human γD-crystallin (HγD),[4] we iterated between experiments and simulations in studying the oxidation-mimicking W42Q mutant as well as the W42R congenital cataract mutant. Despite the difference in charge, the two mutants' thermostabilities and aggregation properties were nearly identical, and both required formation of an internal disulfide bond for aggregation under physiological conditions *in vitro*. All-atom, two-chain tethered Monte Carlo simulations using a knowledge-based potential predicted that, of the six possible internal disulfides within the mutated N-terminal domain, the Cys^{32}–Cys^{41} bond in particular trapped a partially misfolded state prone to specific intermolecular interactions. Experiments confirmed that this bond was indeed required for aggregation *in vitro*. Independently, a recent proteomic study has linked oxidation of these residues to age-onset cataracts *in vivo* (44). Constraining the simulations to the now established disulfide enabled us to propose the most likely structure of the locally unfolded aggregation precursor, as well as a polymerization mechanism that shares key features with domain swapping.

Results

W42R and W42Q Are Similar to Each Other in Stability and Aggregation Properties—Trp^{42}, located in the bottom of the N-terminal domain core (Fig. 1*A*), is particularly sensitive to substitution, as both oxidation-mimicking W42E and W42Q substitutions *in vitro* and the congenital cataract-linked W42R substitution destabilize the protein and lead to aggregation (39, 45, 46). The congenital and oxidation-mimicking substitutions differ in their charge. Nonetheless, the W42R sample showed very similar aggregation properties to W42Q. The purified protein was stable for weeks in storage at 4 °C, but aggregated in a redox-dependent manner when shifted to a higher temperature (Fig. 1*B*). Accommodation of charged residues in protein cores has been the subject of intensive study (47, 48), and thus it was of interest to compare the charged (W42R) and neutral (W42Q) core substitutions. Differential scanning calorimetry (DSC) (Fig. 1*C*) revealed that there was almost no difference in the temperature of the two mutants' unfolding transitions. The available chemical denaturation curves confirm similar stability of W42Q and W42R (39, 46). Table 1 summarizes melting temperatures obtained from DSC of the WT and several mutant constructs.

W42R and W42Q Aggregation Occurs at Solution Redox Potentials Found in Cataractous Lens—Disulfides are known to form in lens cytosol during the course of aging and cataractogenesis as the pool of lens glutathione becomes smaller and more oxidized over time (9, 28, 41, 44). To mimic this physiological oxidation process *in vitro*, we carried out aggregation experiments by first reducing the proteins completely and then placing them at a series of known redox potentials using glutathione buffers. In those experiments, both mutants showed highly oxidation-dependent aggregation at physiological temperature and pH. The range of aggregation-permissive OxD (oxidation degree, defined as the ratio of oxidized/total glutathione) values was comparable with that found in the lens during cataractogenesis (41) (Fig. 1*D*). Although both mutants showed clear aggregation by OxD = 0.18, the WT protein did not show any oxidation-dependent aggregation even at OxD = 0.5, suggesting that aggregation requires a partially unfolded or misfolded conformation accessible at physiological temperatures only as a result of the mutations or comparable damage (Fig. 1*D*).

We have previously reported that the aggregates formed by the temperature-sensitive Trp mutants of γD-crystallin under near physiological conditions are neither disulfide-bridged nor amyloid and expose few hydrophobic residues (39, 45). These observations raised the question of whether a seeding effect exists in this aggregation pathway. Seeding is typical for templated aggregation phenomena, such as amyloidogenesis or crystallization: pre-formed aggregates added to a fresh solution of the protein serve as nuclei, bypassing the lag phase and greatly accelerating the aggregation process. However, as shown in Fig. 2, this is not the case for W42R aggregation under physiological conditions. When a sample from the aggregating protein was mixed in with an identical fresh solution, turbidity initially decreased slightly, and the subsequent lag time was not significantly shorter than that for the initial aggregation. The conformational conversion in W42R HγD does not appear to be a templated one.

Oxidation Causes a Conformational Change in the Mutant Proteins—Intrinsic tryptophan fluorescence is a highly sensitive reporter of conformational change in γD-crystallin, as evidenced by many previous studies (18, 20, 39, 49, 50). As shown in Fig. 3, baseline (non-oxidized) fluorescence spectra show no conformational difference between the WT and mutant proteins. This is consistent with the data in Fig. 1*C*, where unfolding is not detectable for any protein at 37 or 42 °C under reducing conditions. However, when the solution redox potential is oxidizing, the mutants display slightly but reproducibly red-shifted spectra (Fig. 3, *A* and *B*), in contrast with the WT protein (Fig. 3*C*). We attribute this change to increased solvent exposure of at least one Trp residue in the mutant proteins, most likely Trp^{68}, because that is the only Trp residue in the destabilized N-terminal domain of these mutants. This spectral change was observed over the course of 60 min; no red shift was observed when oxidizing agent was omitted (Fig. 3*D*). Thus,

[4] The abbreviations used are: HγD, human γD-crystallin; DSC, differential scanning calorimetry; OxD, oxidation degree; TCEP, tris(2-carboxyethyl) phosphine; RMSD, root mean square deviation.

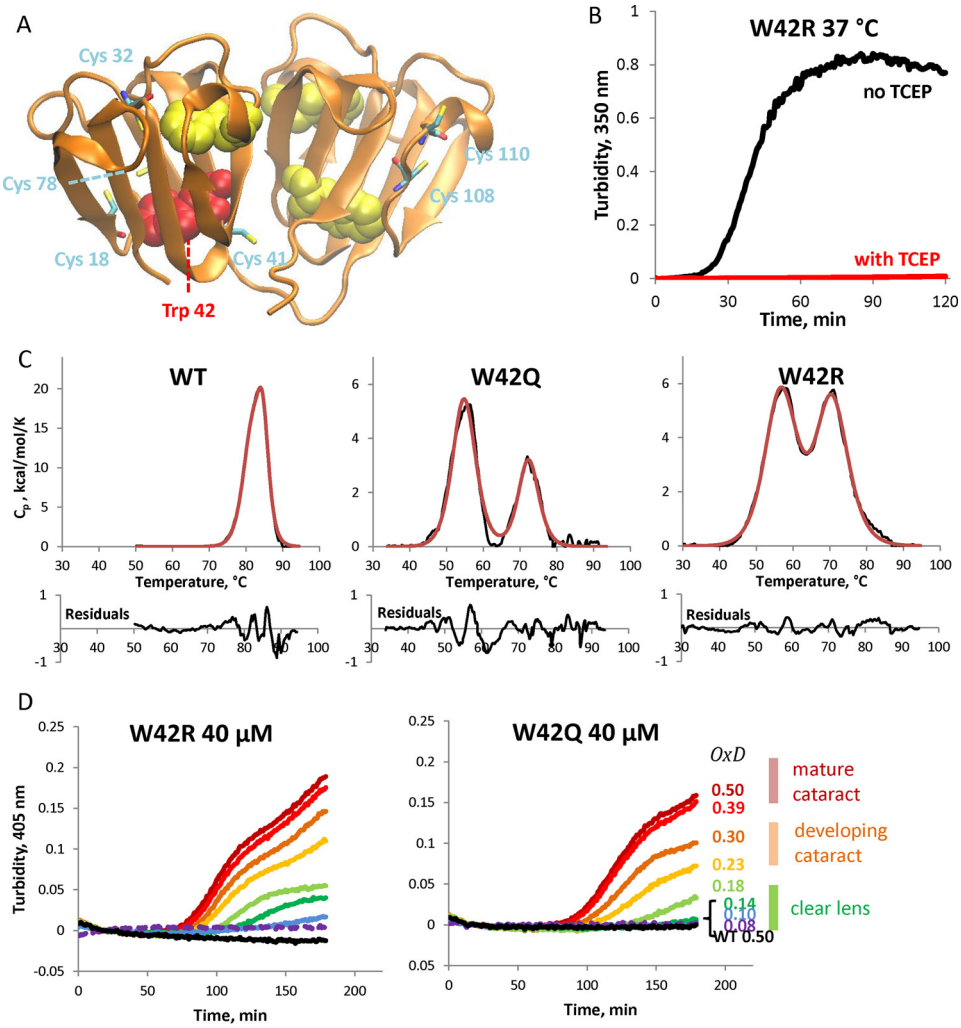


FIGURE 1. **Aggregation of W42Q and W42R human γD-crystallin is highly redox-dependent.** *A*, backbone ribbon diagram of the native, wild-type γD structure (PDB ID 1HK0) showing the Trp residues as space-filling models and Cys residues in stick representation. Trp^{42} is colored *red*. The other Trp residues are shown as *yellow space-filling models*. All Cys residues are also labeled and shown as *stick models*. *B*, turbidity traces of W42R at 150 μM upon shifting from 4 °C to 37 °C in the presence or absence of the reducing agent TCEP. *C*, differential scanning calorimetry unfolding traces (*black*) of WT, W42Q, and W42R constructs with empirical three-state fittings (*red*) reveal that the mutated N-terminal domains have reduced thermostability, but still melt above physiological temperatures. *D*, the redox potential of the solution was set using the glutathione couple such that [GSH] + 2[GSSG] = 2 mM. Proteins were at 40 μM and 37 °C. The legend on the *right* shows the approximate OxD ("oxidation degree") values based on glutathione ratios *in vivo*, calculated from the data in Ref. 41. OxD = 2[GSSG]/(2[GSSG]+[GSH]). For example, OxD = 0.18 corresponds to [GSH]:[GSSG] ratio of 9:1.

TABLE 1

Melting temperatures of the N-terminal and C-terminal domains of the WT and several mutants of HγD

Data are reported as mean ± S.D.; *n* is the number of independent replicates.

Sample	T_m(N)[a]	T_m(C)[b]	*n*
	°C	°C	
WT	85.0 ± 0.4	84.9 ± 0.3	3
W42Q	54.6 ± 0.3	71.1 ± 1.7	2
W42R	56.1 ± 0.7	70.4 ± 0.5	5
C18A/W42Q	51.7 ± 0.2	74.1 ± 0.2	3
C32V/W42Q	60.2 ± 0.2	73.0 ± 0.1	3

[a] Melting temperatures of the N-terminal domain.
[b] Melting temperatures of the C-terminal domain.

aggregation of these mutant proteins correlates with a conformational change stabilized by oxidation.

Simulation of Oxidative Misfolding Reveals Characteristic Early Intermediates—Single chain simulations showed that cataract-associated mutations in the N-terminal domain destabilize the N-terminal domain (Fig. 4*A*), whereas the oxidation-

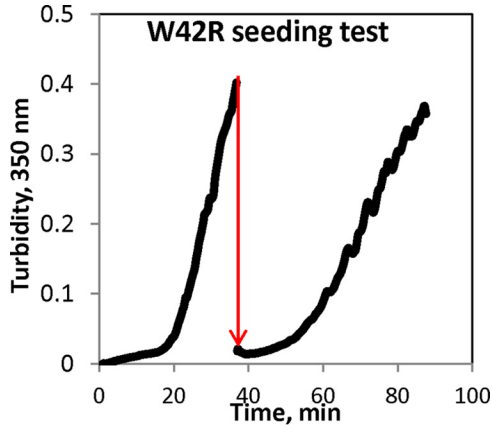


FIGURE 2. **Lack of seeding effect in W42R aggregation 180 μM W42R sample in the middle of its linear range of turbidity development was rapidly mixed 1:10 with fresh 180 μM sample of W42R (the mixing is indicated by the *red arrow*).** No significant change in lag time or aggregation rate was observed upon seeding.

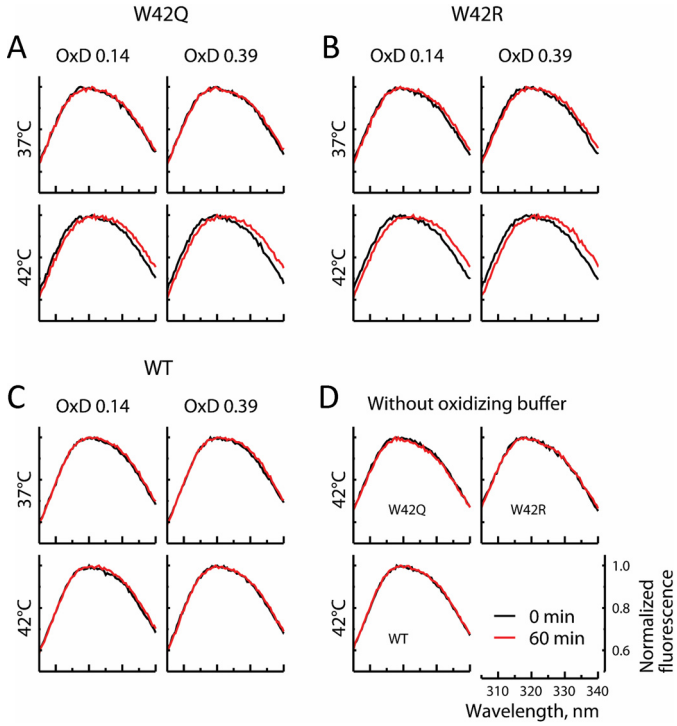


FIGURE 3. **Oxidation causes conformational change in W42Q and W42R.** *A–C*, 4 μM W42Q (*A*), W42R (*B*), and WT (*C*) proteins were incubated for 60 min at 37 or 42 °C and OxD = 0. 14 or 0.39. The protein concentration is below the aggregation range (~10% of that in Fig. 1). Fluorescence traces of the mutant proteins, but not the WT, show a red shift, indicating conformational change, particularly at higher temperature and oxidation. *D*, heating these samples at the same temperatures in the absence of oxidizing agent did not result in any peak shift.

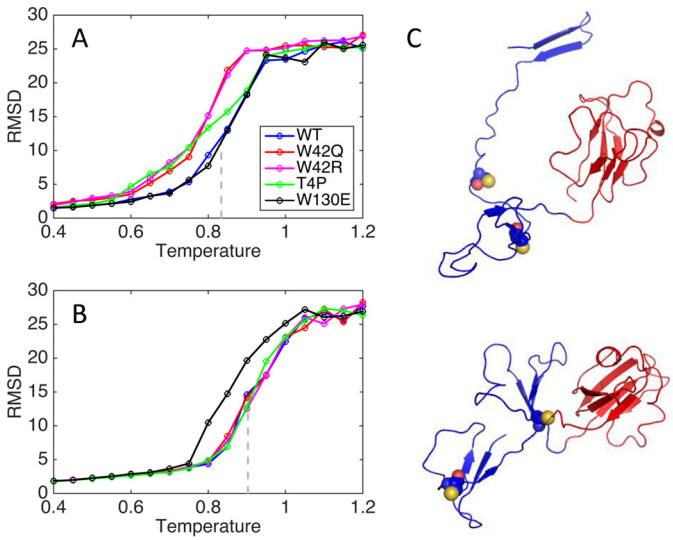


FIGURE 4. **RMSD from native structure for N- and C-terminal domains.** RMSD was averaged over 50 simulations at step 60,000,000. The *gray line* shows the midpoint of the sigmoidal fit to the WT curve. *A*, N-terminal domain. *B*, C-terminal domain. *C*, representative W42R structures populated at $T = 0.800$ with N-terminal domains in *blue* and C-terminal domains in *red*. Cys^{32} and Cys^{41} are shown as space-filling models.

mimicking mutation W130E in the C-terminal domain destabilizes the C-terminal domain relative to WT (Fig. 4*B*). In particular, W42Q and W42R both led to destabilization of the N-terminal domain, with very similar apparent melting curves, consistent with the experiments (Fig. 1*C*). Even in the WT protein, the N-terminal domain was less stable than the C-terminal domain, also consistent with past experimental results (18, 50, 51) and simulations (52). Also consistent with the experiments, the W130E mutation largely eliminated the differential domain stability, thus leading to a more cooperative unfolding transition (45). Partially unfolded structures from W42R simulations at intermediate temperature show that separation of the two N-terminal Greek keys and detachment of the N-terminal hairpin are early events in the unfolding process (Fig. 3*C*). As we discuss below, the flexibility of the loop connecting the detached hairpin to the rest of the domain core may allow Cys residues that are distant in the native structure to come together and form a disulfide bond.

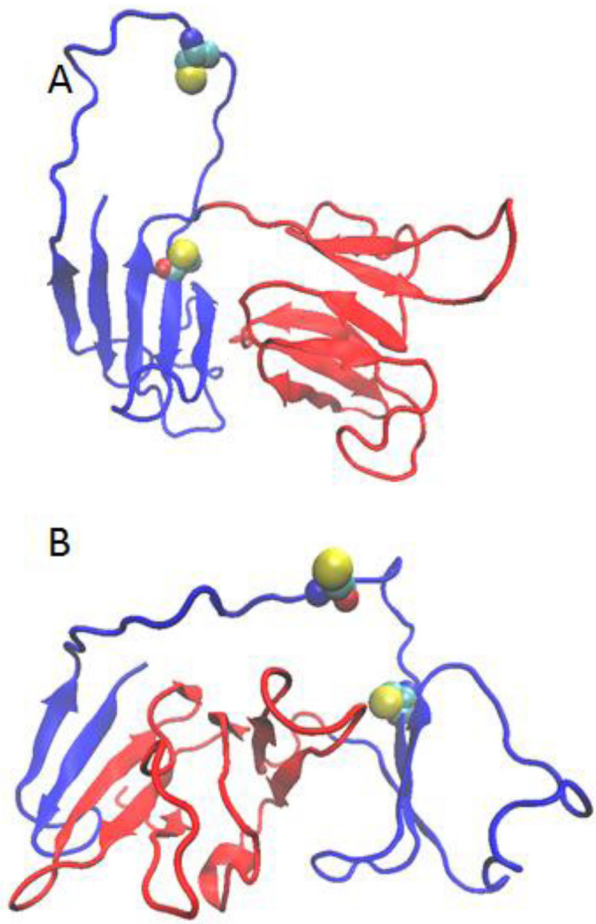


FIGURE 5. **Examples of internally rearranged W42R structures.** The N-terminal domain is *blue*; the C-terminal domain is *red*. A, rearrangement to enable binding of strand 1 to strand 8 within the N-terminal domain. *B*, rearrangement to enable binding of strand 1 to strand 14. Structures at simulated $T = 0.8$. Cys^{32} and Cys^{41} are shown as space-filling models.

Of 300 independently simulated unfolding pathways, five were remarkable in that the N-terminal hairpin, having detached from its native position, annealed to another β-sheet on the molecule, extending it. The final states of three of these runs are shown in Fig. 5*A*, and the other two are shown in Fig. 5*B*. In both cases, the observed internal topological rearrangements suggest possible binding sites for aggregate propagation.

Two-chain Tethered Simulations Predict Internal Disulfide Structures Prone to Extensive Intermolecular Interactions—Whole-protein mass spectrometry experiments (39) revealed that most monomers within the aggregated fraction

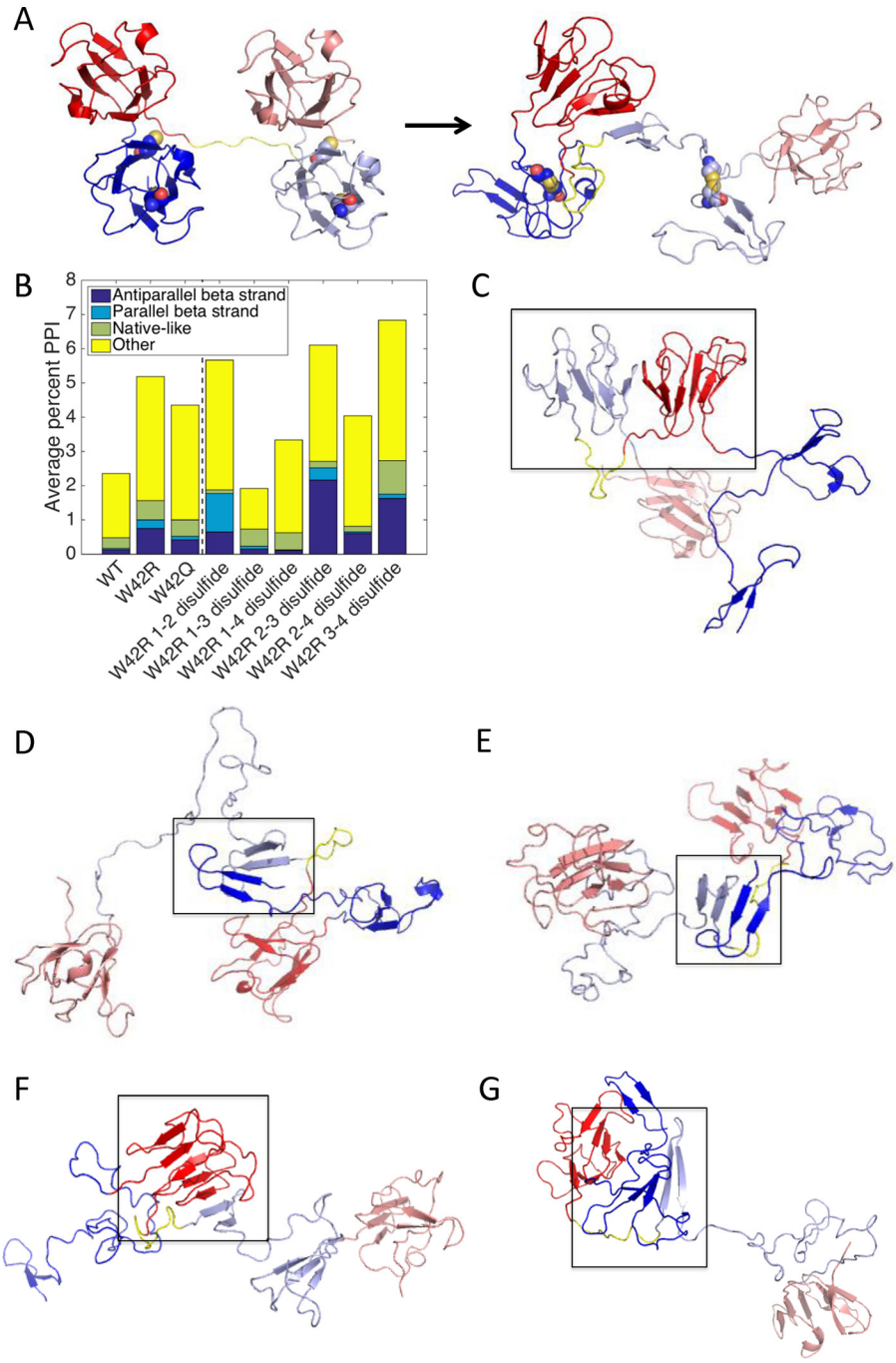
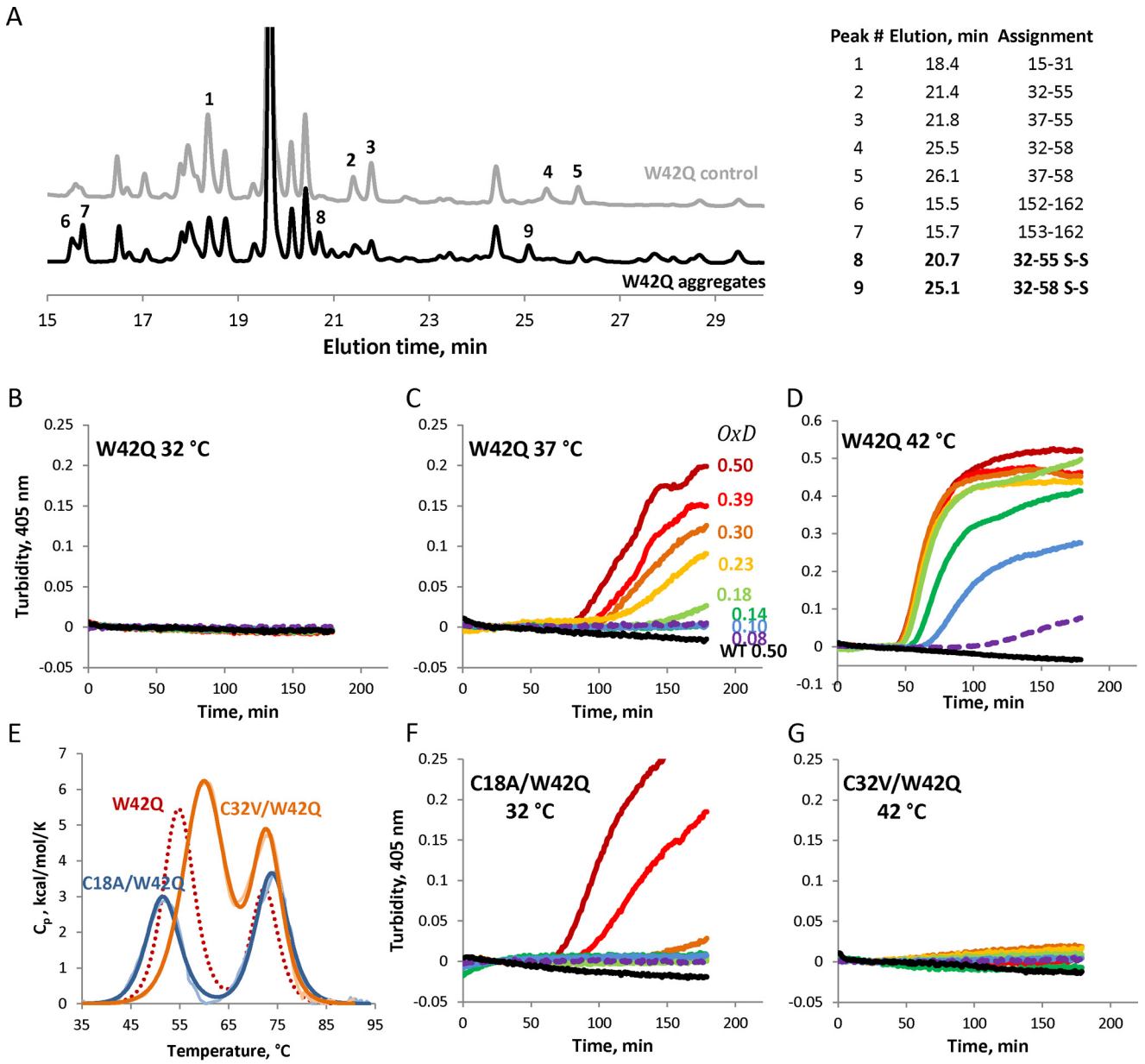


FIGURE 6. **Protein-protein interactions from two-molecule simulations of γD crystallin and predicted structure of aggregates.** *A*, initial simulation structure and final frame from a simulation trajectory, showing the N-terminal hairpin of one molecule binding to the folded C-terminal domain of the other molecule via interactions between strand 1 of the first molecule and the second strand of the C-terminal Greek key of the second molecule. *B*, the percentage of simulation frames showing protein-protein interactions (*PPI*), averaged over 300 trajectories. *C–G*, representative classes of interacting structures. *C*, a native-like interaction, typical of the 1–4 disulfide. *D*, a parallel strand 1-strand 1 interaction, typical of the 1–2 disulfide. *E*, an antiparallel strand 1-strand 1 interaction, common for 1–2 and 2–3 disulfide. *F*, an antiparallel strand 1-strand 14 interaction, typical of the 2–3 disulfide. *G*, a representative structure from the "Other" category. N-terminal domains are *blue*; C-terminal domains are *red*. The simulated flexible linker between the two protein molecules is in *yellow*.

contain only one disulfide bond. HγD contains six Cys residues (Fig. 1*A*), so we set out to determine whether a particular internal disulfide is most likely to generate intermolecular interactions that could lead to a polymeric state. This was accomplished by simulating two identical mutant chains tethered by a flexible $(Gly\text{-}Ser)_6$ linker. Disulfide bonding was simulated by placing a spring potential between pairs of thiols, leading to a preferred 2-Å final sulfur-sulfur distance. In all cases, the simulated internal disulfides formed prior to the intermolecular interactions (Fig. 6*A*). Protein-protein interactions were found in a relatively small fraction of simulation frames for WT and mutant proteins, including W42R with internal disulfide bonds (Fig. 6*B*). Interestingly, disulfides between cysteine residues that are adjacent in sequence led to a greater number of protein-protein interactions than disulfides between non-adjacent cysteines (Fig. 6*B*), consis-



Peak #	Elution, min	Assignment
1	18.4	15-31
2	21.4	32-55
3	21.8	37-55
4	25.5	32-58
5	26.1	37-58
6	15.5	152-162
7	15.7	153-162
8	**20.7**	**32-55 S-S**
9	**25.1**	**32-58 S-S**

FIGURE 7. **Cys^{32}–Cys^{41} is the predominant internal disulfide enabling W42Q aggregation.** *A*, HPLC traces of aggregated (OxD = 0.2) and control (non-oxidized) W42Q samples following trypsin digestion under non-reducing conditions. Assignments for selected peaks, using electrospray mass spectrometry, are shown at the *right*. *B–D*, oxidative aggregation of W42Q is highly temperature-dependent in assays carried out as in Fig. 1. *E*, Cys mutations in the W42Q N-terminal domain alter its thermostability. Fitting for the W42Q DSC unfolding shown in *red* (*dashed line*) from Fig. 1*C* is shown for comparison. *F* and *G*, temperatures for the OxD aggregation assays on the C18A/W42Q (*F*) and C32V/W42Q (*G*) constructs were chosen to control for the differential thermostability. The former mutant aggregates already at 32 °C, whereas the latter does not do so even at 42 °C.

tent with previous work on lattice proteins (53). The type of interaction (Fig. 6, *C–G*) depended on the particular mutant and disulfide bond formed. In particular, disulfide bonding between the second and third cysteines in the N-terminal domain led to the greatest number of intermolecular antiparallel β-sheet interactions. Therefore, this disulfide bond, between residues Cys^{32} and Cys^{41}, was predicted as the likeliest to generate redox-sensitive aggregates.

LC/MS and Mutagenesis Identify Cys^{32}–Cys^{41} as the Predominant Disulfide Cross-link Responsible for Aggregation—We mapped the predominant disulfide in W42Q aggregates produced by mild oxidation (OxD = 0.2) using mass spectrometry. When samples were digested under non-reducing conditions, two new peaks appeared in the HPLC traces of the aggregated sample relative to the stock control (Fig. 7*A*). The identities of select HPLC peaks were assigned by electrospray mass spectrometry, where disulfide-containing peptides are distinguished by a −2-Da mass shift. These experiments revealed Cys^{32}–Cys^{41} as the likely cross-link, which is in agreement with both our computational results and the most recent lens proteomic data (44). To detect the disulfide directly, we carried out tryptic digestion without a reduction step, as reported previously (54). A drawback of this approach is increased likelihood of missed tryptic cuts, as well as chymotryptic cleavages. However, these features also offered the advantage of multiple independent detections of disulfide-bonded peaks within the same

experiment. Furthermore, tryptic digestion is known to be particularly inefficient within tight secondary structures such as β-turns. Accordingly, most of the digested samples had a missed cleavage between the bonded Cys residues, at Arg^{36}. Digestion efficiency at this position was also higher in the absence of the disulfide bond (Fig. 7*A*). Because missed cleavage at Arg^{36} makes Cys^{32} and Cys^{41} part of the same peptide, we were able to conclude that the disulfide identified here is, in fact, internal, rather than intermolecular. Because size-exclusion traces of the OxD = 0.2 supernatant did not reveal any dimers, we conclude that intermolecular disulfides were unlikely to form under these conditions.

Thus, the predominant disulfide was identified in W42Q. Unfortunately, disulfides in W42R could not be mapped by LC/MS in the same way because the mutation introduces an additional tryptic cut site, resulting in short, hydrophilic peptides that could not be easily separated by HPLC. However, given the previous results, we felt it was fair to assume that the two mutations are essentially interchangeable and have treated them as such in interpreting this study.

We further confirmed the disulfide mapping results using mutagenesis. Cys residues were mutated in the background of W42Q. In this background, Ala substitutions at Cys^{41} or Cys^{78} did not yield soluble protein, likely due to excessive loss of stability of the already destabilized mutant N-terminal domain. The C18A/W42Q double mutant could be purified, albeit with reduced yield. The C32A/W42Q construct's yield was impracticably low, so we resorted instead to a Val substitution, which is present at this position in many related crystallin sequences. These two mutants are predicted to have divergent aggregation phenotypes: C18A/W42Q is predicted to retain its ability to aggregate by the pathway we have proposed, whereas C32V/W42Q is expected to lack this ability.

Oxidative aggregation was found to be highly temperature-dependent, consistent with a requirement of conformational change (39, 45). Thus, W42Q showed no aggregation at 32 °C even under the most oxidizing conditions employed; oxidation-dependent aggregation at 37 °C; and strong aggregation even at low OxD values at 42 °C (Fig. 7, B–D). As shown in Fig. 7*E*, the C32V/W42Q double mutant is ~5.6 °C more stable than W42Q, whereas C18A/W42Q is ~2.9 °C less stable. To control for the effect of thermostability, we carried out oxidative aggregation assays at 32 °C for C18A/W42Q and 42 °C for C32V/W42Q. In contrast to W42Q, C18A/W42Q aggregated in an oxidation-dependent manner already at 32 °C, whereas C32V/W42Q showed no aggregation even at 42 °C (Fig. 6, *F* and *G*).

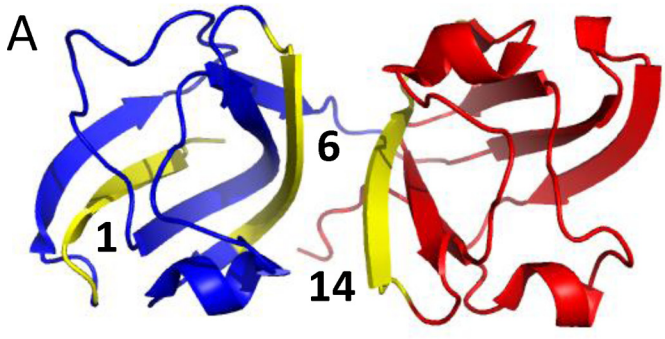
Molecules with Cys^{32}–Cys^{41} Internal Disulfide Are Predicted to Aggregate by β-Sheet Completion—Having established experimentally by LC/MS, and confirmed by mutagenesis, that the Cys^{32}–Cys^{41} internal disulfide was indeed required for aggregation, we carried out a more detailed analysis of two-chain simulations involving this disulfide bond. The observed parallel and antiparallel interactions were classified based on which β-strands were in contact. Most of these interactions were between strand 1 and strand 14, or between the first strands of the two proteins (Fig. 8; see also Fig. 6, *E* and F). In the former case, the interaction could be perpetuated, as the N-terminal hairpin of each protein binds to the folded or mostly folded domain of the next protein, extending a β-sheet that is formed in the native structure (model in Fig. 9).

Discussion

Age-onset cataract is a highly prevalent disease as well as a case study in protein aging. Its pathology has multiple contributing factors, including exposure to UV light, tobacco, and certain specific chemicals. However, the effect is generally the same: aggregation of the eye lens crystallins (5, 13, 16, 17, 55, 56). In line with the multiple causality of the disease, several distinct biochemical perturbations *in vitro* have been shown to cause crystallin aggregation (49, 57–59). Cataract-associated point mutations, especially ones mimicking *in vivo* chemical modifications, have been highly useful in dissecting the mechanism of aggregation (21, 39, 46) and developing interventions to modify it (60). Synergistic effects are likely among the various modes of modification or damage. Thus, our combined computational and experimental results suggest a link between Trp oxidation, often the result of UV exposure, and disulfide bonding, a marker of more general oxidative stress in the aging lens.

Our all-atom Monte Carlo program has been used to successfully predict folded states and relative mutant stabilities of single proteins in previous publications (61, 62). Here, we used the program to simulate interactions between two proteins by connecting them by a flexible linker. Such an approach is commonly used in experimental single-molecule pulling studies, including of HγD (63), and has been used in the past to computationally study protein-protein interactions (64, 65). Unlike the approach introduced by Dima and Thirumalai (66), we did not constrain the distance between proteins' centers of mass; however, both approaches can in principle be used to simulate protein concentration. Although we are currently developing a version of the Monte Carlo program with multiple freely diffusing chains, the peptide linker approach is a straightforward way to adapt a program designed for single chain simulations to study of the initial steps in protein aggregation. In fact, the approach yields reasonable results: the strands most often involved in protein-protein interaction are protected in the native structure but become exposed early in single-chain simulations. We expect that peptide-based drugs targeting these strands may slow or prevent aggregation of γ-crystallin molecules.

Using molecular dynamics simulations with subsequent annealing of two γ-crystallin molecules, Das *et al.* (67) found interactions between the unfolded N-terminal domain and strands 13–15. Similar regions interact in our simulations, but the interactions are more specific. Thus, we observe β-sheet completion, involving hydrogen bonding between β-strands. The speed of our program relative to conventional molecular dynamics techniques allows for a greater sampling of configurational space necessary to obtain such structures, as well as obviates the need to simulate under highly denaturing conditions (such as the 8 M urea used by Das *et al.* (67)). We propose that aggregation proceeds from oligomerization in which the N-terminal hairpin binds to an exposed region of the C-terminal domain. Specifically, the N-terminal hairpin detaches from the N-terminal domain, and strand 1 forms antiparallel hydrogen-bonding interactions with strand 14. Although such a



strand	residues			strand	residues		
1	1	-	9	9	85	-	96
2	10	-	19	10	97	-	106
3	20	-	30	11	107	-	118
4	31	-	39	12	119	-	127
5	40	-	49	13	128	-	137
6	50	-	59	14	138	-	146
7	60	-	74	15	147	-	161
8	75	-	84	16	162	-	173

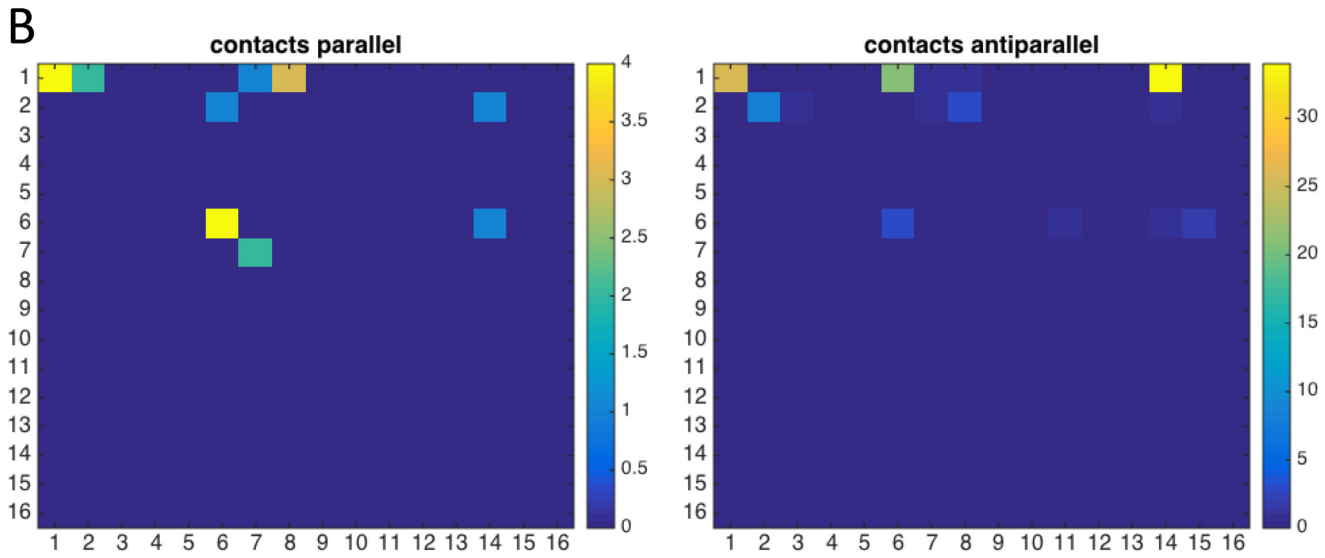


FIGURE 8. **Strand-strand contact maps from tethered two-chain W42R simulations involving the 2–3 disulfide.** *A*, the most commonly interacting β-strands are strand 1 and the two strands most involved in the domain interface (highlighted in *yellow*). The N-terminal domain is *blue*; the C-terminal domain is *red*. *B*, combined contact maps for all simulations showing frequency of pairwise non-native strand-strand interactions.

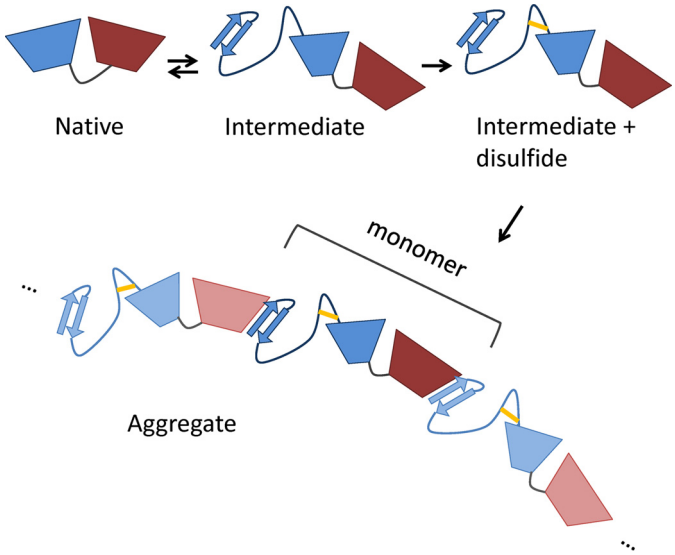


FIGURE 9. **A graphical model of aggregation by β-sheet completion between strand 1, in the N-terminal β-hairpin extruded due to formation of the Cys^{32}–Cys^{41} internal disulfide bridge, and strand 14 in the C-terminal domain.**

structure remains to be experimentally validated, it is consistent with the lower folding stability of the N-terminal domain relative to the C-terminal domain and the observation that many cataract-associated mutations in γ-crystallins cluster in or near the N-terminal hairpin (20). It also agrees with the recent report that the N-terminal β-hairpin is involved in domain swapping in HγD polyproteins under mechanical force (63).

That internal disulfide bonds stabilize specific partially unfolded states in HγD was not appreciated until recently (39). Although our statistical potential does not fully account for the distinct chemical properties (*e.g.* hydrophobicity) of disulfides and thiols, we have incorporated the geometric constraints on folding that result from disulfide formation. We found that disulfides between cysteines close together in sequence, particularly Cys^{32} and Cys^{41}, stabilized a partially unfolded intermediate state prone to specific intermolecular interactions. It is worth noting that disulfide bonding is critical to the aggregation pathways of other well known disease-related proteins, including superoxide dismutase 1 and β_2-microglobulin (34, 35, 37, 38). Conversely, disulfide engineering has been used to trap important folding intermediates in a variety of proteins (68). More generally, covalent linkage of residues that are nearby in the primary structure is known to promote misfolding and polymerization. This has been shown both in theoretical lattice models of proteins (53) and, for instance, in experimental (69) and computational (70) studies of the aggregation of $A\beta_{1-40}$ with the Asp^{23}–Lys^{28} lactam bridge.

Surprisingly, among the possible disulfides in HγD, our study pointed to one in particular, Cys^{32}–Cys^{41}, as most likely to stabilize the aggregation precursor. Our results agree with a recent proteomic study linking the oxidation state of Cys^{32} and Cys^{41} in γD-crystallin to cataractous aggregation in human lens (44). Thus, it may be possible to reconstitute the *in vivo* cataractogenic biochemical pathway both *in vitro* and *in silico* to reveal structures of transient intermediates and specific misfolded states en route to "amorphous" aggregation. It is noteworthy that oxidation of Cys^{41} to a disulfide is one of the strongest correlates of human (but not murine) cataract development identified by Monnier and colleagues (44), and that it is found in both γD-crystallin and γC-crystallin. However, it has been unclear in the proteomic studies whether such disulfides are intra- or intermolecular. The aggregation in our *in vitro* assays is due to an internal bond, because both Cys residues are part of the same disulfide-containing peptide (Fig. 7*A*). Studying aggregates by high-resolution structural techniques *in vitro* is very challenging, leading to a paucity of atomic-level mechanistic detail. Simulated structures based on the Cys^{32}–Cys^{41} bridge provide a possible molecular mechanism for misfolding and polymerization into light-scattering aggregates (Fig. 9).

The aggregation model proposed in Fig. 9 shares many features with classical domain swapping, as proposed by the Eisenberg group (71), yet the polymers do not reconstitute the native state of the monomeric protein. Horwich (1) pointed out the possibility that domain swapping may reconstitute not the native state, but a partially unfolded or misfolded state that is at least kinetically stable. Because we have observed intramolecular strand 1-strand 14 binding in single-chain simulations (Fig. 5), this may be an example of such a mechanism. Stability of the internally rearranged monomer depends on the length and flexibility of the linker connecting it to the rest of the domain core. Our simulations have shown that this linker is sufficiently long and flexible in the absence of a disulfide bond. Shortening or straining a linker between two structural elements is the most common way to generate classical domain swapping (72, 73). Formation of the Cys^{32}–Cys^{41} disulfide, which we find to be required for polymerization, creates a loop within this linker sequence and may thus accomplish both. Reducing this disulfide bond is expected to make the monomeric state favorable again, consistent with our observation that the addition of reducing agent can reverse the aggregation process (39).

We note that the internal rearrangement to form a "cyclic" monomer is expected to compete with the aggregation process, yet it is also likely to facilitate formation of the internal disulfide bond needed for aggregation. A potentially better way to inhibit aggregation would be to block the binding edge of strand 14 with a peptide or small-molecule inhibitor that would compete with the N-terminal β-hairpin. One biological precedent for protein/peptide binding via β-sheet completion is found in the binding of PDZ domains to their ligands (74). Finally, a major challenge for inhibition of classical domain-swapping reactions is that any intervention to stabilize the native state is likely to also stabilize the polymer. If a partially misfolded state is the origin of the swapping, however, then inhibitors can be targeted to a specific part of the native structure.

The primary model of protein-protein interactions suggested by our simulations extends the self-recognition and domain-swapping paradigms currently associated with protein aggregation. Misfolding simulations of multidomain proteins such as titin suggest a role for domain-swapping and/or self-similar amyloid-like interactions in misfolding and aggregation (75, 76), and these studies have now been extended to HγD (63). In our predicted oligomerizing structure, strand 14 is sandwiched between two homologous β-strands, one belonging to the native structure (strand 13) and the other belonging to another domain or protein (strand 1). Thus, the non-native extended β-sheet structure is formed via native-like interactions that are not actually present in the native structure but that mimic native contacts on the other side of the β-strand. It will be interesting to see whether such "native-like" interactions between regions homologous to natively interacting ones are found in other aggregating proteins.

Experimental Procedures

Protein Expression and Purification—All proteins were expressed and purified as described (39). Briefly, BL21-RIL *Escherichia coli* (Agilent) were transformed with pET16b plasmids containing the relevant cDNAs under the T7 promoter with no epitope tags. Cells were induced at 18 °C overnight. Cell pellets were stored at −80 °C. Ammonium sulfate was added slowly to cleared cell lysate to ~30% to precipitate less soluble contaminants, and then crystallins were precipitated by increasing ammonium sulfate concentration to ~50%. The pelleted proteins were washed and further purified by size-exclusion chromatography. Purity was >95% by SDS-PAGE. Typical yields for the W42Q and W42R mutants were 30–40 mg/liter cell culture in Super Broth, 10–15 mg/liter for C18A/W42Q, and 50–70 mg/liter for C32V/W42Q and WT proteins.

Differential Scanning Calorimetry—Thermal denaturation of WT and mutant proteins was carried out by differential scanning calorimetry (Nano DSC, TA Instruments) using 0.3 or 0.6 mg of protein. Prior to the experiment, proteins were reduced with TCEP for 1–2 h at room temperature, and then dialyzed three times against DSC buffer (10 mM sodium phosphate, 50 mM NaCl, pH 7). TCEP to a final concentration of 0.5 mM was added to the samples immediately prior to DSC. The buffer from the dialysis reservoir was used as reference for the DSC experiments. The samples were heated from 25 to 90 °C at a scan rate of 1 °C/h. Thermodynamic parameters were derived by fitting the data to a sum of two 2-state unfolding models, one for each domain, using NanoAnalyze (TA Instruments).

Aggregation Assays—Aggregation of purified samples was induced by shifting to 37 °C or as indicated in sample buffer (10 mM ammonium acetate, 50 mM NaCl, pH 7) with 1 mM EDTA. Turbidity was monitored using a FLUOstar Optima plate reader (BMG Labtech) and polypropylene 96-well plates (Greiner) in 200-μl volume with no agitation. In the case of Figs. 1*B* and 2, the measurements were performed in a quartz cuvette in a spectrophotometer. Oxidation-dependent aggregation experiments were performed in the same way but with a controlled degree of oxidation. Purified protein samples were reduced by incubation with 0.5 mM TCEP for 30–60 min at room temperature with gentle agitation. The reducing agent

was removed by 2–3 rounds of dialysis against 4 liters of sample buffer, one of which was overnight. Degree of oxidation (OxD) was set using mixtures of reduced and oxidized glutathione (GSH and GSSG) at a total concentration of 2 mM (counting GSSG as two molecules), according to the formula

$$\mathrm{OxD} = \frac{2[\mathrm{GSSG}]}{[\mathrm{GSH}] + 2[\mathrm{GSSG}]} \quad \text{(Eq. 1)}$$

Fluorescence—Protein samples (4 μM, in 10 mM ammonium acetate buffer, 50 mM NaCl, 1 mM EDTA, pH 7) were equilibrated in heated cuvettes at the indicated temperatures for 10 min prior to the experiment. Intrinsic tryptophan fluorescence spectra were obtained using a Cary Eclipse fluorimeter (Varian) with excitation at 280 nm and emission scanned between 300 and 400 nm to make sure the signal reached baseline by 400 nm.

Monte Carlo Unfolding Simulations—All-atom (non-hydrogen) Monte Carlo simulations were performed using a program described in previous publications, developed by the Shakhnovich group (61, 62, 77, 78). The program incorporates a knowledge-based potential, with terms for contact energy, hydrogen-bonding, torsional angle, and side-chain torsional terms, as well as a term describing relative orientations of aromatic residues. The move set consists of rotations about backbone and side-chain torsional angles, with bonds and angles held fixed. Software for the Monte Carlo Protein Unfolding (MCPU) program is available on the Shakhnovich group website.

Human γD crystallin (Protein Data Bank (PDB) ID 1HK0) was used as the initial structure in simulations. Mutations (W42Q, W42R, T4P, and W130E) were introduced using PyMOL. For WT and each mutant, an initial simulation was run at low temperature ($T = 0.150$ in simulation units) for 2,000,000 steps. Starting from the final structure of this simulation, a 60,000,000-step simulation was run, for each of 32 temperatures, ranging from $T = 0.1$–3.2. The final frame was extracted from each of 50 separate runs, and the average RMSD from the native structure was calculated for the N-terminal and C-terminal domains separately. Curves representing the dependence of RMSD from the native structure on temperature were fit to a sigmoidal function; because the domains were fitted separately, the sigmoid corresponded to a two-state model of unfolding. Representative structures were obtained for W42R by clustering contact maps (residue-residue α-carbon distance less than 10 Å) for final simulation frames, using the MATLAB clusterdata function with a cutoff of 0.9. Images were generated for representative structures from the two most populated clusters. A further 300 simulations at $T = 0.8$ were then conducted.

Two-chain Tethered Unfolding Simulations—Two-molecule single chain simulations were carried out for two γ-crystallin molecules connected by a 12-residue linker. The amino acid sequence of the linker was GSGSGSGSGSGS, and the initial structure is shown in Fig. 5*A*. Each simulation was 80,000,000 Monte Carlo steps in length. 300 simulations were run for each of WT, W42R, and W42Q. In addition, simulations were run with a harmonic constraint applied to each pair of disulfides within the N-terminal domain, holding sulfur atoms of the cysteine residues ~2 Å apart. The same potential maintained the disulfide once formed. Cys-Cys angles were neglected for these simulations, unless disallowed by steric effects.

Simulations were mined for protein-protein interactions every 5,000,000 frames. Residues were said to be in contact if their α-carbons were within 10 Å of each other. A protein-protein interaction was counted when more than 50 pairwise residue-residue contacts were detected between proteins. An interaction was said to be an antiparallel β-strand interaction if it included more than six residues in a row in an antiparallel configuration. An interaction was said to be a parallel β-strand interaction if the interaction included more than six residues in a row in a parallel configuration. The interaction was said to be native-like if greater than 10 of the interactions were present between domains of the native structure (see, for instance, Fig. 5*C*). For parallel and antiparallel β-strand interactions, the β-strands involved in the interaction were recorded, according to the numbering system in Fig. 7.

Disulfide Mapping—The predominant disulfide bond was mapped in W42Q aggregates produced upon incubation at 37 °C and OxD = 0.2 for ~2 h in sample buffer with 1 mM EDTA. Aggregates were pelleted by centrifugation of the turbid mixture at 14,500 × *g* for 5–10 min. The pelleted fraction was washed three times with sample buffer, and then resuspended in 0.1 M MES buffer, pH 5.5. The resuspended sample was kept at 37 °C for 20–30 min to allow dissociation of weakly bound protein, leaving stable aggregates. These aggregates were solubilized by the addition of guanidine hydrochloride to a final concentration of 4 M and denatured by incubating at 37 °C for 3 h, at pH 5 and in the presence of 1 mM EDTA. The low pH and chelator were used to inhibit Cys oxidation and disulfide scrambling during denaturation. Denatured samples were refolded by gradual dilution of the guanidine-containing buffer with dilution buffer (0.1 M MES, pH 5.5, 1 mM EDTA) at the rate of 25% dilution every 15 min. Once [guanidine] had been brought down to 2.6 M, samples were centrifuged and concentration was determined by NanoDrop (Thermo Scientific). Dilution buffer was then added to reduce [guanidine] to 2.0 M, and then L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK)-treated bovine trypsin (Sigma-Aldrich) was added at the ratio of 1:20 w/w. Dilution continued until [guanidine] fell to 1.0 M. Samples were centrifuged again, and concentration was determined, and additional trypsin was added to bring the total trypsin:protein ratio to 1:10. Samples were then incubated for 12–16 h at 37 °C to allow digestion. Afterward, digests were centrifuged to remove any precipitated material, and trypsin was activity quenched by the addition of trifluoroacetic acid to 0.3% v/v final concentration. The high trypsin:protein ratio was necessitated by the partly denaturing, non-reduced digestion conditions. Stock (pre-reduced) W42Q sample was used as a control for the denaturation and digestion steps to check for adventitious disulfides that might arise during sample handling. We used W42Q in preference to W42R to facilitate isolation of Cys^{41}-containing peptides from the tryptic digests.

Tryptic digests were separated on a 5-μm, 250 × 4.6-mm C18 column (Higgins Analytics) on an Agilent 1200 HPLC instrument using a 36-min 5–60% gradient of acetonitrile with 0.1% trifluoroacetic acid. Absorbance at 280 nm was monitored to enable quantitative comparisons of peak intensity indepen-

dent of their ionization efficiency. Peaks of interest were collected manually and identified separately by brief HPLC over a 2.7-μm C18 column (Agilent) coupled to an Agilent 6230 electrospray ionization mass spectrometer. Multiply charged states were identified and deconvoluted manually.

Author Contributions—E. S., J. C. W., J. A. K., and E. I. S. designed the research; J. C. W. carried out the computational modeling; E. S. carried out the aggregation and mutagenesis experiments; E. S. and B. V. A. carried out calorimetry and fluorescence experiments; E. S. and M. S. carried out LC/MS experiments; E. S. and J. C. W. wrote the manuscript; B. V. A., M. S., J. A. K., and E. I. S. edited the manuscript; J. A. K. and E. I. S. supervised the research. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments—We are grateful to Prof. Elizabeth Nolan (Massachusetts Institute of Technology) for advice on experimental design; to Cameron Haase-Pettingell for protein expression; to Phoom Chairatana and Lisa S. Cunden for training on HPLC and electrospray ionization-mass spectrometry (ESI-MS); and to Dr. Catherine Goodman for editorial suggestions.

References

1. Horwich, A. (2002) Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J. Clin. Invest.* **110,** 1221–1232
2. Chiti, F., and Dobson, C. M. (2009) Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5,** 15–22
3. Bartels, T., Choi, J. G., and Selkoe, D. J. (2011) α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* **477,** 107–110
4. Aguzzi, A., and Calella, A. M. (2009) Prions: protein aggregation and infectious diseases. *Physiol. Rev.* **89,** 1105–1152
5. Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N. H., Slingsby, C., and Tardieu, A. (2004) Ageing and vision: structure, stability and function of lens crystallins. *Prog. Biophys. Mol. Biol.* **86,** 407–485
6. Uversky, V. N. (2010) Mysterious oligomerization of the amyloidogenic proteins. *FEBS J.* **277,** 2940–2953
7. Borgia, M. B., Nickson, A. A., Clarke, J., and Hounslow, M. J. (2013) A mechanistic model for amorphous protein aggregation of immunoglobulin-like domains. *J. Am. Chem. Soc.* **135,** 6456–6464
8. Rakhit, R., Crow, J. P., Lepock, J. R., Kondejewski, L. H., Cashman, N. R., and Chakrabartty, A. (2004) Monomeric Cu,Zn-superoxide dismutase is a common misfolding intermediate in the oxidation models of sporadic and familial amyotrophic lateral sclerosis. *J. Biol. Chem.* **279,** 15499–15504
9. Truscott, R. J. W. (2005) Age-related nuclear cataract: oxidation is the key. *Exp. Eye Res.* **80,** 709–725
10. Fujiwara, H., Hasegawa, M., Dohmae, N., Kawashima, A., Masliah, E., Goldberg, M. S., Shen, J., Takio, K., and Iwatsubo, T. (2002) α-Synuclein is phosphorylated in synucleinopathy lesions. *Nat. Cell Biol.* **4,** 160–164
11. de Graff, A. M. R., Hazoglou, M. J., and Dill, K. A. (2016) Highly charged proteins: the Achilles' heel of aging proteomes. *Structure* **24,** 329–336
12. Kennedy, S. R., Loeb, L. A., and Herr, A. J. (2012) Somatic mutations in aging, cancer and neurodegeneration. *Mech. Ageing Dev.* **133,** 118–126
13. Michael, R., and Bron, A. J. (2011) The ageing lens and cataract: a model of normal and pathological ageing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366,** 1278–1292
14. Wride, M. A. (2011) Lens fibre cell differentiation and organelle loss: many paths lead to clarity. *Philos. Trans. R. Soc. Lond. B Biol. Sci* **366,** 1219–1233
15. Shang, F., and Taylor, A. (2004) Function of the ubiquitin proteolytic pathway in the eye. *Exp. Eye Res.* **78,** 1–14
16. Moreau, K. L., and King, J. A. (2012) Protein misfolding and aggregation in cataract disease and prospects for prevention. *Trends Mol. Med.* **18,** 273-282
17. Colon, W., and Kelly, J. W. (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation *in vitro*. *Biochemistry* **31,** 8654–8660
18. Flaugh, S. L., Mills, I. A., and King, J. (2006) Glutamine deamidation destabilizes human γD-crystallin and lowers the kinetic barrier to unfolding. *J. Biol. Chem.* **281,** 30782–30793
19. Mahler, B., Doddapaneni, K., Kleckner, I., Yuan, C., Wistow, G., and Wu, Z. (2011) Characterization of a transient unfolding intermediate in a core mutant of γS-crystallin. *J. Mol. Biol.* **405,** 840–850
20. Serebryany, E., and King, J. A. (2014) The βγ-crystallins: Native state stability and pathways to aggregation. *Prog. Biophys. Mol. Biol.* **115,** 32–41
21. Takata, T., Oxford, J. T., Demeler, B., and Lampi, K. J. (2008) Deamidation destabilizes and triggers aggregation of a lens protein, βA3-crystallin. *Protein Sci.* **17,** 1565–1575
22. Metlapally, S., Costello, M. J., Gilliland, K. O., Ramamurthy, B., Krishna, P. V., Balasubramanian, D., and Johnsen, S. (2008) Analysis of nuclear fiber cell cytoplasmic texture in advanced cataractous lenses from Indian subjects using Debye-Bueche theory. *Exp. Eye Res.* **86,** 434–444
23. Truscott, R. J. W. (2007) Eye lens proteins and cataracts. in *Protein Misfolding, Aggregation, and Conformational Diseases* (Uversky V. N., and Fink, A. L., eds), pp. 435–447, Springer, New York
24. Sandilands, A., Hutcheson, A. M., Long, H. A., Prescott, A. R., Vrensen, G., Löster, J., Klopp, N., Lutz, R. B., Graw, J., Masaki, S., Dobson, C. M., MacPhee, C. E., and Quinlan, R. A. (2002) Altered aggregation properties of mutant γ-crystallins cause inherited cataract. *EMBO J.* **21,** 6005–6014
25. Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J., and Benedek, G. B. (2001) Crystal cataracts: human genetic cataract caused by protein crystallization. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 6116–6120
26. Hains, P. G., and Truscott, R. J. W. (2007) Post-translational modifications in the nuclear region of young, aged, and cataract human lenses. *J. Proteome Res.* **6,** 3935–3943
27. Hains, P. G., and Truscott, R. J. W. (2008) Proteomic analysis of the oxidation of cysteine residues in human age-related nuclear cataract lenses. *Biochim. Biophys. Acta* **1784,** 1959–1964
28. Spector, A., and Roy, D. (1978) Disulfide-linked high molecular-weight protein associated with human cataract. *Proc. Natl. Acad. Sci. U.S.A.* **75,** 3244–3248
29. Lampi, K. J., Wilmarth, P. A., Murray, M. R., and David, L. L. (2014) Lens β-crystallins: the role of deamidation and related modifications in aging and cataract. *Prog. Biophys. Mol. Biol.* **115,** 21–31
30. Hua, Q. X., Jia, W., Frank, B. H., Phillips, N. F. B., and Weiss, M. A. (2002) A protein caught in a kinetic trap: structures and stabilities of insulin disulfide isomers. *Biochemistry* **41,** 14700–14715
31. Chang, J. Y., Li, L., and Lai, P. H. (2001) A major kinetic trap for the oxidative folding of human epidermal growth factor. *J. Biol. Chem.* **276,** 4845–4852
32. Arolas, J. L., Aviles, F. X., Chang, J. Y., and Ventura, S. (2006) Folding of small disulfide-rich proteins: clarifying the puzzle. *Trends Biochem. Sci.* **31,** 292–301
33. Fraga, H., Bech-Serra, J.-J., Canals, F., Ortega, G., Millet, O., and Ventura, S. (2014) The mitochondrial intermembrane space oxireductase Mia40 funnels the oxidative folding pathway of the cytochrome *c* oxidase assembly protein Cox19. *J. Biol. Chem.* **289,** 9852–9864
34. Toichi, K., Yamanaka, K., and Furukawa, Y. (2013) Disulfide scrambling describes the oligomer formation of superoxide dismutase (SOD1) proteins in the familial form of amyotrophic lateral sclerosis. *J. Biol. Chem.* **288,** 4970–4980
35. Ding, F., and Dokholyan, N. V. (2008) Dynamical roles of metal ions and the disulfide bond in Cu,Zn superoxide dismutase folding and aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 19696–19701
36. Karamanos, T. K., Kalverda, A. P., Thompson, G. S., and Radford, S. E. (2014) Visualization of transient protein-protein interactions that promote or inhibit amyloid assembly. *Mol. Cell* **55,** 214–226
37. Chen, Y., and Dokholyan, N. V. (2005) A single disulfide bond differentiates aggregation pathways of β_2-microglobulin. *J. Mol. Biol.* **354,** 473–482
38. Smith, D. P., and Radford, S. E. (2001) Role of the single disulphide bond of β_2-microglobulin in amyloidosis *in vitro*. *Protein Sci.* **10,** 1775–1784
39. Serebryany, E., and King, J. A. (2015) Wild-type human γD-crystallin pro-

motes aggregation of its oxidation-mimicking, misfolding-prone W42Q mutant. *J. Biol. Chem.* **290,** 11491–11503
40. Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., Rodnina, M. V., and Komar, A. A. (2016) Synonymous codons direct cotranslational folding toward different protein conformations. *Mol. Cell* **61,** 341–351
41. Friedburg, D., and Manthey, K. F. (1973) Glutathione and NADP linked enzymes in human senile cataract. *Exp. Eye Res.* **15,** 173–177
42. Ozaki, Y., Mizuno, A., Itoh, K., and Iriyama, K. (1987) Inter- and intramolecular disulfide bond formation and related structural changes in the lens proteins: a Raman spectroscopic study *in vivo* of lens aging. *J. Biol. Chem.* **262,** 15545–15551
43. Hains, P. G., and Truscott, R. J. W. (2010) Age-dependent deamidation of lifelong proteins in the human lens. *Invest. Ophthalmol. Vis. Sci.* **51,** 3107–3114
44. Fan, X., Zhou, S., Wang, B., Hom, G., Guo, M., Li, B., Yang, J., Vaysburg, D., and Monnier, V. M. (2015) Evidence of highly conserved β-crystallin disulfidome that can be mimicked by *in vitro* oxidation in age-related human cataract and glutathione depleted mouse lens. *Mol. Cell. Proteomics* **14,** 3211–3223
45. Serebryany, E., Takata, T., Erickson, E., Schafheimer, N., Wang, Y., and King, J. A. (2016) Aggregation of Trp > Glu point mutants of human γ-D crystallin provides a model for hereditary or UV-induced cataract. *Protein Sci.* **25,** 1115–1128
46. Ji, F., Jung, J., Koharudin, L. M. I., and Gronenborn, A. M. (2013) The human W42R γD-crystallin mutant structure provides a link between congenital and age-related cataracts. *J. Biol. Chem.* **288,** 99–109
47. Isom, D. G., Castañeda, C. A., Cannon, B. R., Velu, P. D., and García-Moreno E, B. (2010) Charges in the hydrophobic interior of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 16096–16100
48. Robinson, A. C., Castañeda, C. A., Schlessman, J. L., and García-Moreno E, B. (2014) Structural and thermodynamic consequences of burial of an artificial ion pair in the hydrophobic interior of a protein. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 11685–11690
49. Kosinski-Collins, M. S., and King, J. (2003) *In vitro* unfolding, refolding, and polymerization of human γD crystallin, a protein involved in cataract formation. *Protein Sci.* **12,** 480–490
50. Mills, I. A., Flaugh, S. L., Kosinski-Collins, M. S., and King, J. A. (2007) Folding and stability of the isolated Greek key domains of the long-lived human lens proteins γD-crystallin and γS-crystallin. *Protein Sci.* **16,** 2427–2444
51. Flaugh, S. L., Kosinski-Collins, M. S., and King, J. (2005) Interdomain side-chain interactions in human γD crystallin influencing folding and stability. *Protein Sci.* **14,** 2030–2043
52. Das, P., King, J. A., and Zhou, R. (2010) β-Strand interactions at the domain interface critical for the stability of human lens γD-crystallin. *Protein Sci.* **19,** 131–140
53. Abkevich, V. I., and Shakhnovich, E. I. (2000) What can disulfide bonds tell us about protein energetics, function and folding: Simulations and bioinformatics analysis. *J. Mol. Biol.* **300,** 975–985
54. Wommack, A. J., Ziarek, J. J., Tomaras, J., Chileveru, H. R., Zhang, Y., Wagner, G., and Nolan, E. M. (2014) Discovery and characterization of a disulfide-locked C_2-symmetric defensin peptide. *J. Am. Chem. Soc.* **136,** 13494–13497
55. Zhang, J., Yan, H., Löfgren, S., Tian, X., and Lou, M. F. (2012) Ultraviolet radiation-induced cataract in mice: the effect of age and the potential biochemical mechanism. *Invest. Ophthalmol. Vis. Sci.* **53,** 7276–7285
56. Robman, L., and Taylor, H. (2005) External factors in the development of cataract. *Eye* **19,** 1074–1082
57. Moran, S. D., Zhang, T. O., Decatur, S. M., and Zanni, M. T. (2013) Amyloid fiber formation in human γD-crystallin induced by UV-B photodamage. *Biochemistry* **52,** 6169–6181
58. Schafheimer, N., and King, J. (2013) Tryptophan cluster protects human D-crystallin from ultraviolet radiation-induced photoaggregation *in vitro*. *Photochem. Photobiol.* **89,** 1106–1115
59. Quintanar, L., Domínguez-Calva, J. A., Serebryany, E., Rivillas-Acevedo, L., Haase-Pettingell, C., Amero, C., and King, J. A. (2016) Copper and zinc ions specifically promote nonamyloid aggregation of the highly stable human γ-D crystallin. *ACS Chem. Biol.* **11,** 263–272
60. 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., *et al.* (2015) Lanosterol reverses protein aggregation in cataracts. *Nature* **523,** 607–611
61. Yang, J. S., Chen, W. W., Skolnick, J., and Shakhnovich, E. I. (2007) All-atom ab initio folding of a diverse set of proteins. *Structure* **15,** 53–63
62. Tian, J., Woodard, J. C., Whitney, A., and Shakhnovich, E. I. (2015) Thermal stabilization of dihydrofolate reductase using Monte Carlo unfolding simulations and its functional consequences. *PLoS Comput. Biol.* **11,** e1004207
63. Garcia-Manyes, S., Giganti, D., Badilla, C. L., Lezamiz, A., Perales-Calvo, J., Beedle, A. E. M., and Fernández, J. M. (2016) Single-molecule force spectroscopy predicts a misfolded, domain-swapped conformation in human D-crystallin protein. *J. Biol. Chem.* **291,** 4226–4235
64. Levy, Y., Cho, S. S., Onuchic, J. N., and Wolynes, P. G. (2005) A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *J. Mol. Biol.* **346,** 1121–1145
65. Levy, Y., Wolynes, P. G., and Onuchic, J. N. (2004) Protein topology determines binding mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 511–516
66. Dima, R. I., and Thirumalai, D. (2002) Exploring protein aggregation and self-propagation using lattice models: Phase diagram and kinetics. *Protein Sci.* **11,** 1036–1049
67. Das, P., King, J. A., and Zhou, R. (2011) Aggregation of γ-crystallins associated with human cataracts via domain swapping at the C-terminal β-strands. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 10514–10519
68. Narayan, M. (2012) Disulfide bonds: protein folding and subcellular protein trafficking. *FEBS J.* **279,** 2272–2282
69. Sciarretta, K. L., Gordon, D. J., Petkova, A. T., Tycko, R., and Meredith, S. C. (2005) Aβ40-Lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry* **44,** 6003–6014
70. Reddy, G., Straub, J. E., and Thirumalai, D. (2009) Influence of preformed Asp23-Lys28 salt bridge on the conformational fluctuations of monomers and dimers of Aβ peptides with implications for rates of fibril formation. *J. Phys. Chem. B* **113,** 1162–1172
71. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Domain swapping: entangling alliances between proteins. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 3127–3131
72. Rousseau, F., Schymkowitz, J. W. H., and Itzhaki, L. S. (2003) The unfolding story of three-dimensional domain swapping. *Structure* **11,** 243–251
73. Cutler, T. A., Mills, B. M., Lubin, D. J., Chong, L. T., and Loh, S. N. (2009) Effect of interdomain linker length on an antagonistic folding-unfolding equilibrium between two protein domains. *J. Mol. Biol.* **386,** 854–868
74. Maisonneuve, P., Caillet-Saguy, C., Vaney M-C, Edoo B.-Z., Sawyer, K., Raynal, B., Delepierre, M., Lafon, M., Cordier, F., and Wolff, N. (2016) Molecular basis of the interaction of the human protein tyrosine phosphatase non-receptor type 4 (PTPN4) with the mitogen-activated protein kinase p38γ. *J. Biol. Chem.* **291,** 16699–16708
75. Zheng, W., Schafer, N. P., and Wolynes, P. G. (2013) Frustration in the energy landscapes of multidomain protein misfolding. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 1680–1685
76. Borgia, A., Kemplen, K. R., Borgia, M. B., Soranno, A., Shammas, S., Wunderlich, B., Nettels, D., Best, R. B., Clarke, J., and Schuler, B. (2015) Transient misfolding dominates multidomain protein folding. *Nat. Commun.* **6,** 8861
77. Yang, J. S., Wallin, S., and Shakhnovich, E. I. (2008) Universality and diversity of folding mechanics for three-helix bundle proteins. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 895–900
78. Xu, J., Huang, L., and Shakhnovich, E. I. (2011) The ensemble folding kinetics of the FBP28 WW domain revealed by an all-atom Monte Carlo simulation in a knowledge-based potential. *Proteins* **79,** 1704–1714