

**Title:** Accumulative deamidation of human lens protein  $\gamma$ S-crystallin leads to partially unfolded intermediates with enhanced aggregation propensity

**Authors (Names and Addresses):** Calvin J. Vetter, David C. Thorn, Samuel G. Wheeler, Charlie Mundorff, Kate Halverson, John A. Carver, Larry L. David, Kirsten J. Lampi

**Corresponding Author:** Kirsten J. Lampi – [lampik@ohsu.edu](mailto:lampik@ohsu.edu)

### Abstract:

Age-related cataract is a major cause of blindness worldwide. Yet, the molecular mechanisms whereby large, light scattering aggregates form is poorly understood, because of the complexity of the aggregates isolated from human lenses. The predominant proteins in the lens are structural proteins called crystallins. The  $\gamma$ S-crystallin is heavily modified in cataractous lenses by deamidation, which introduces a negative charge at labile asparagine residues. The effects of deamidation at asparagines, N14, N76, and N143, were mimicked by replacing the asparagine with aspartate using site-directed mutagenesis. The effects of these surface deamidations on the stability, unfolding, and aggregation properties of  $\gamma$ S were determined using dynamic light scattering, chemical and thermal-denaturation, and hydrogen-deuterium exchange with mass spectrometry. We found that a small population of all the deamidation mimics aggregated directly into large light scattering bodies with a radius greater than 10 nm that contributed 14-60% of the total scattering intensity compared to 7% for WT under the same conditions. A possible mechanism was identified under partially denaturing conditions, where deamidation led to significantly more rapid unfolding and aggregation particularly for N76D compared to WT. The triple mutant was further destabilized, reflecting the enhanced aggregation properties of N14D and N143D. Thus, the effects of deamidation were both site-specific and cumulative.  $\alpha$ A-crystallin was ineffective at acting as a chaperone to prevent the aggregation of destabilized, deamidated  $\gamma$ S. It is concluded that surface deamidations, while causing minimal structural disruption individually, progressively destabilize crystallin proteins, leading to their unfolding and precipitation in aged and cataractous lenses.

### Abbreviations List:

Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), dithiothreitol (DTT), sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), guanidine hydrochloride (GuHCl), Dynamic, static, and multi-angle light scattering (DLS, SLS, and MALS), hydrogen deuterium exchange (H/D), mass spectrometry (MS), wild type (WT), triple mutant (TM), N-terminal domain (N-td), C-terminal domain (C-td).

### Introduction

In mammals, the eye lens plays a crucial role in vision via transmitting, refracting and focusing light onto the retina. Lens functionality is maintained via a high concentration of crystallin proteins (up to 300-500 mg/mL in the center of the lens) that are arranged in a supramolecular array with short-range order (1). The lens is a unique organ in that it has no blood supply and there is no protein turnover in its fiber cells because they lack the organelles for protein synthesis and degradation. As a result, crystallins are long-lived proteins (2).

Cataract arises from opacification and concomitant light scattering associated with the unfolding, aggregation and precipitation of crystallins. Cataract is the major cause of blindness in the world and is the leading cause of low vision in the United States (3). Age-related cataract is the most prevalent type although many crystallin mutations give rise to congenital and early-onset cataract. The molecular mechanism(s) whereby cumulative modifications of crystallins lead to the formation of large, light scattering aggregates in the aging lens is poorly understood, because of the complexity of the insoluble aggregates isolated from human lenses (4-9).

There are three types of mammalian crystallins,  $\alpha$ ,  $\beta$  and  $\gamma$ , which are all comprised of a variety of isoforms (10, 11). The two  $\alpha$ -crystallins (A and B) are small heat-shock molecular chaperone proteins. The  $\beta$ - and  $\gamma$ -crystallins are homologous to each other and are not related to the  $\alpha$ -crystallins. The  $\beta$ - and  $\gamma$ -crystallins are  $\beta$ -sheet-rich, two-domain proteins with the  $\beta$ -sheets adopting a Greek key arrangement. The Greek key motif contains four  $\beta$ -strands. In  $\gamma$ S-crystallin ( $\gamma$ S), each domain comprises two Greek key motifs, i.e. a total of eight  $\beta$ -strands (Fig. 1 – the NMR structure of  $\gamma$ S). The domains in the  $\gamma$ -crystallins are monomers with their domains linked by a connecting peptide that is bent, whereas in  $\beta$ -crystallins the connecting peptide is either bent or extended leading to oligomer formation with hydrophobic interactions stabilizing the interface between the domains (12-15).

Due to the absence of protein turnover, the lens is an incubator that accumulates multiple crystallin modifications over the individual's lifetime. Deamidation is a major age-related crystallin modification associated with insolubilization of crystallins *in vivo* (6, 16, 17). Deamidation sites at buried interfaces between  $\beta$ - and  $\gamma$ -crystallin domains, while disruptive to protein stability, are not the most prevalent deamidation sites found *in vivo*. The most prevalent are in solvent-exposed regions (6, 18, 19). *In vivo*,  $\gamma$ S undergoes significant deamidation with age, in this case the conversion of asparagine residues, principally at surface residues (N14, N76 and N143), each to four possible isomers of aspartic acid, i.e. L-Asp, L-isoAsp, D-Asp and D-isoAsp (20-22). The location of these deamidation sites are indicated on the solution NMR-derived structure of  $\gamma$ S (Fig. 1) (23). At these labile surface Asn residues, levels range from 25-60% at N14, 50-80% at N76, and 20-90% at N143 in  $\gamma$ S implying a strong association between these deamidation sites and cataract formation (6, 18, 19, 22, 24, 25). The extent of deamidation in the entire crystallin mixture, including these three  $\gamma$ S sites (particularly N76 and N143), is significantly greater in cataract lenses compared to age-matched, non-cataract controls, and significantly greater in the insoluble proteins than the soluble proteins from these lenses, implying that these modifications contribute to cataract formation (6, 18, 22, 24, 25). It is expected that the other isomers would affect the protein structure even more.

Previous studies have investigated the effects of deamidation of  $\gamma$ S by mimicking deamidated residues with L-Asp, replacing L-Asn via mutagenesis and introducing a negative charge at the site of modification (7, 26). We recently reported that single-point deamidations on the surface of  $\gamma$ S can alter its conformational dynamics, providing a potential mechanism for the unfolding and insolubilization of deamidated crystallins in the eye lens (27). In this present study, the effects of deamidation (conversion to L-Asp) at each these three sites, along with the triple mutant (N14D/N76D/N143D, TM), was investigated to determine the accumulative effects of the surface deamidations on the stability, unfolding, and aggregation properties of  $\gamma$ S. We found that a small population of all the deamidation mimics aggregated into large light scattering bodies. A

possible mechanism was identified under partially denaturing conditions, where deamidation led to more rapid unfolding and aggregation particularly for N76D compared to WT. The triple mutant was further destabilized, reflecting the enhanced aggregation properties of N14D and N143D. Thus, the effects of deamidation were both site-specific and cumulative. Furthermore,  $\alpha$ A-crystallin was ineffective at acting as a chaperone to prevent the aggregation of destabilized, deamidated  $\gamma$ S. It is concluded that surface deamidations, while causing minimal structural disruption individually, progressively destabilize crystallin proteins, leading to their unfolding and precipitation in aged and cataractous lenses.

## Experimental

**Recombinant expression and purification of human  $\gamma$ S and deamidation mimics.** The plasmid containing gene for  $\gamma$ S wild-type was obtained from the plasmid repository at DNASU (<https://dnasu.org/DNASU/Home.do>). Single deamidation mimics of N14, N76, N143, plus the triple mimic of all three sites were generated via site-directed mutagenesis using the QuikChangeXL Kit (Agilent Technologies, CA). Protein was expressed in a pE-SUMO (Small Ubiquitin-like Modifier) vector containing a N-terminal 6XHis tagged fusion protein (LifeSensors Inc., PA) using *Escherichia coli*. Culture of cells took place at 37 °C, protein expression was induced with IPTG and growth continued at 37 °C for 4 h. Purification of the protein used HisPur Cobalt Resin following the manufacturer's protocol (Thermo Fisher Scientific, Waltham, MA.). The removal of the His Tag and SUMO was performed using Ulp-1 protease (Ubl-specific protease 1). Protein was then run through HisPur Cobalt Resin again to remove the tag. The protein was aliquoted, 2.4 mM dithiothreitol (DTT) was added and samples stored at -80 °C. Typical yields of protein were several mg per 100 mL of *E. coli* culture. The purity and masses of all expressed proteins were confirmed by mass spectrometry and gel electrophoresis to match predicted masses and to be greater than 95% pure. Protein concentration was determined by measuring absorbance at 280 nm and calculating concentrations using the protein extinction coefficient calculated using the ExPasy ProtParam tool (human  $\gamma$ S 42,860 cm<sup>2</sup> M<sup>-1</sup>).

**Dynamic light scattering (DLS) of WT and deamidated  $\gamma$ S.** Proteins were concentrated to 1 or 5 mg/mL using centrifugation (Amicon Ultra-15 Filter Units (10 kDa MWCO), Millipore Sigma, Burlington, MA). Samples were exchanged into Buffer A (29 mM Na<sub>2</sub>HPO<sub>4</sub>, 29 mM NaH<sub>2</sub>PO<sub>4</sub>, 100 mM KCl, 1 mM EDTA and 1 mM DTT) using dialysis and cleared of large particulate matter using 0.1  $\mu$ m pore size filters (Ultrafree-MC Centrifugal Filter Units, Millipore Sigma). Samples of 30  $\mu$ L were loaded into a 384 well plate in replicates of 5 for each  $\gamma$ S and the DLS measured at 25 °C using a Dynapro Plate Reader III (Wyatt Technology Corp., Santa Barbara, CA). The time-dependent fluctuations of scattered light intensity were measured and the translational diffusion coefficient ( $D_t$ ) was calculated by the autocorrelation function of the DYNAMICS software (Wyatt Tech.). The  $D_t$  was used to calculate the hydrodynamic radii ( $R_h$ ) (20). The percent of DLS intensity was plotted for the distributions of different size species in an intensity-weighted regulation graph. Sizes were distributed into different peaks of 1-10 nm, 10-100 nm, and 100-1000 nm and the percent polydispersity (%Pd) was calculated as the distribution of radii within each peak. Weight average molar masses estimated from DLS ( $M_w$ -R) were predicted from the  $R_h$  and assuming a sphere by DYNAMICS software.

### ***Size-exclusion chromatography and multi-angle light scattering (SEC-MALS) of WT and deamidated $\gamma$ S.***

WT and deamidated  $\gamma$ S were buffer exchanged and concentrated as described for DLS above. The  $M_w$  were determined from the measured Rayleigh light scattering (Dawn Heleos II and ASTRA VII Software, Wyatt Tech.). A dn/dc of 0.193 and absorbance at 280 nm or a refractive index detector were used to determine concentration (Optilab REX, Wyatt Tech.) (7). Fractions eluting under each peak were analyzed by SDS-PAGE. The eluting fractions containing the monomer peaks for WT and TM  $\gamma$ S were next collected, concentrated as above for DLS and incubated for 6 days at 37 °C. Samples were then reanalyzed by SEC-MALS to detect aggregates that had formed from the monomers with time. The polydispersity index (PDI) was determined from the ratio of  $M_w$  to  $M_n$  (i.e.  $M_w/M_n$ ), where  $M_n$  is the weight average number of molecules at each time interval. A monodisperse sample would have a PDI of 1.000. All values were calculated in ASTRA VII software and variability is a percentage of the average value under the protein peak (Wyatt Tech.). The percent of protein eluting in each peak was determined from the amount of protein recovered for that peak compared to the total amount of protein recovered in all the peaks using the in-line refractive index detector.

***Equilibrium chemical unfolding of WT and deamidated  $\gamma$ S.*** Proteins were denatured with increasing amounts of guanidine hydrochloride (GuHCl) in buffer containing 100 mM NaPi (pH 7.0), 1.1 mM EDTA, and 5.5 mM DTT at 37°C overnight with a final protein concentration of 10  $\mu$ g/mL. Samples were excited at 295 nm and emission spectra measured from 310-400 nm with slit widths set at 0.5 mm using a spectrofluoropolarimeter (Photon Technology Inc., NJ). The ratio of the 360 nm to 320 nm fluorescence intensities were used to calculate percent protein unfolded at each GuHCl concentration. Non-linear line fit was performed and data was best fitted to a sigmoidal two state unfolding curve using PRISM 6 (GraphPad Software, CA). The Gibbs free energy,  $\Delta G_D^0$  (kcal/mol), of unfolding was determined as previously described using the  $RT \ln K_D$  values calculated for  $K_D = f_D/f_N$  from the transition region of the GuHCl denaturation curves (28). The change in Gibbs free energy,  $\Delta \Delta G_D^0$  (kcal/mol), of unfolding was calculated using the difference between the  $\Delta G_D^0$  of WT (7.58 kcal/mol), and the  $\Delta G_D^0$  of each deamidated  $\gamma$ S.

***Global hydrogen-deuterium (H/D) exchange of WT and deamidated  $\gamma$ S.*** Proteins were incubated in 2.75 M GuHCl and pulse labeled for 1 min in D<sub>2</sub>O buffer. The number of amide hydrogens exchanged with deuterium was determined by measurement of the masses of whole proteins. Samples were diluted to 0.9 mg/mL in buffer containing 50 mM NaPi (pH 7.4), 300 mM NaCl, 10 mM imidazole, and 10 mM DTE. Diluted protein (10  $\mu$ L) was mixed with 40  $\mu$ L of 3.44 M GuHCl in H<sub>2</sub>O exchange buffer (20 mM NaPi (pH 7.4), 2.5 mM TCEP, and 1 mM EDTA). After unfolding for 0-32 min at 30 °C, 10  $\mu$ L of each GuHCl-treated sample was mixed with 90  $\mu$ L of D<sub>2</sub>O containing exchange buffer for 1 min at 30 °C. Labeling was quenched by addition of 25  $\mu$ L ice-cold 0.8 M phosphate buffer (pH 2.4) and samples immediately frozen on dry ice. Masses of unexchanged proteins were determined by similar treatment, but by adding 90  $\mu$ L of H<sub>2</sub>O exchange buffer instead. Mass analysis of  $\gamma$ S was performed by individually and reproducibly thawing each sample and immediate injection onto an ice emerged 1 x 8 mm Opti-Trap™ protein trap cartridge (Optimized Technologies, Oregon City, OR) at 50  $\mu$ L/min in a mobile phase containing 0.1% formic acid and 2% acetonitrile. After washing for 3 min, proteins were eluted in mobile phase containing 0.1% formic acid and 80% acetonitrile and mass measured using a LTQ Velos ion trap mass spectrometer (Thermo Fisher Sci.) using MS scans in

enhanced mode. Spectra acquired during the approximately 30 sec protein elution were averaged, and measurements at each time point repeated 3 times. The region of the spectra from m/z 906-918 containing the +23 ion was used to calculate the number of deuterium exchanged into the various species at each time during GuHCl unfolding. The area of the 3 major peaks of deuterated protein were calculated and expressed as the percent of total area of all observed peaks.

Samples were also exchanged into D<sub>2</sub>O buffer as described above but without denaturant. The amount of deuterium uptake with time was measured by mass spectrometry using the methods above.

***Chemically-induced aggregation of WT and deamidated  $\gamma$ S.*** Unfolding and aggregation induced by GuHCl were monitored using a BioTek Synergy 2 microplate reader (BioTek, Winooski, VT). Proteins were initially concentrated to 10 mg/mL using Ultra-15 Filter Units (10 kDa MWCO) and filtered with Millex-GV Syringe Filter Units (0.22  $\mu$ m pore size). Samples were then prepared at 5 mg/mL in 25 mM NaPi (pH 7.2), 100 mM NaCl, with 10 mM DTT and 2 M GuHCl. Protein samples (60  $\mu$ l) were transferred to black  $\mu$ CLEAR® 96-well microplates (Greiner Bio-One, Monroe, NC; Catalogue No. 675096) which were then sealed with MiniStrips™ (Excel Scientific, Victorville, CA; Catalogue No. SP-2x8-50). Microplates were incubated with continuous shaking (1140 rpm) at 37 °C. Turbidity (405 nm), thioflavin T fluorescence (excitation/emission at 440 $\pm$ 20/485 $\pm$ 10 nm) and intrinsic fluorescence (excitation/emission at 284 $\pm$ 5/360 $\pm$ 20 nm) were measured every minute for 12 hours.

Aggregation kinetics were fitted to a single Boltzmann function using Origin 9.0 (OriginLab Corporation, Northampton, MA) and the fitting parameters were used to calculate the maximum rate of aggregation as previously described (29). Statistical significance was assessed using a two-sample t-test with unequal variances.

***Thermally-induced aggregation of WT and deamidated  $\gamma$ S.*** Proteins were heated at 70 °C (i.e. the approximate midpoint of thermal unfolding(30)) and changes in turbidity measured. An aliquot of 100  $\mu$ L of each crystallin was diluted to 24  $\mu$ M using 100 mM NaPi (pH 7.4), 100 mM KCl, 1 mM EDTA, 2.5 mM TCEP and heated to 70 °C in Peltier Thermal Cycler-200 (MJ Research, CA). Turbidity of each solution was measured in 96-well plate as the change in optical density at 405 nm using Synergy-HT Microplate Reader (BioTek). Initial rates of aggregation were determined by linear regression using PRISM 6 (GraphPad Software) and p-values determined using t-tests.

For chaperone assays, either WT or TM were mixed with  $\alpha$ A-crystallin ( $\alpha$ A) in a ratio of 2:1,  $\gamma$ S: $\alpha$ A, to a final volume of 100  $\mu$ L containing a total of 100  $\mu$ g of protein. Samples were diluted as needed with 100 mM NaPi (pH 7.4), 100 mM KCl, 1 mM EDTA, and 2.5 mM TCEP. Triplicates of each sample were prepared and incubated at 65 °C. Turbidity at 405 nm was measured on a plate reader at 0, 5, 10, 15, 30 and 60 minutes. Control samples contained 66.6  $\mu$ g of WT or TM in 100  $\mu$ L to match the amount of  $\gamma$ S protein in experimental samples. To determine the identity of the precipitating species, samples were centrifuged, and the soluble proteins separated from the pellet. The pellet was washed, recentrifuged and the remaining insoluble proteins resolubilized. Both the soluble and insoluble proteins were analyzed by SDS-



PAGE. After visualization with Coomassie blue staining, protein bands were excised and the protein identified by mass spectrometry.

## Results

### ***Surface deamidations increased protein aggregation of human $\gamma$ S under native conditions.***

Dynamic light scattering indicated that deamidation of  $\gamma$ S increased its propensity to form large aggregates (Fig. 2). Freshly prepared WT  $\gamma$ S at a concentration of 1 mg/mL, showed a predominant single peak with a mean  $R_h$  of 2.5 nm (1.5-4.9 nm range) and a smaller amount of scatter coming from a species between 10-100 nm in the intensity-weighted size distribution (Fig. 2 and Table 1). This is consistent with literature values of 2.0-2.5 nm for monomeric  $\gamma$ S (30, 31). At a higher concentration of 5 mg/mL, similar  $R_h$  was detected for the main peak, but with a greater DLS intensity plus an additional species >100 nm (Fig. 2 and Table 1).

In contrast to WT  $\gamma$ S, the deamidation mimics showed predominantly two to three distinct peaks with a greater percent of intensity coming from the second and third peaks and less coming from the predicted monomer peak, peak 1. Peak 1 of the deamidated mimics had a  $R_h$  ranging from 2.7 to 3.3 nm (Table 1); while not significantly different from the WT, the N14D and N76D mimics' Peak 1 had a wider size distribution, ranging from 1.5-9.9 nm and significantly greater polydispersity (%Pd) (35-37% compared to 19% for WT). Peak 2, with a  $R_h$  ranging from 10-100 nm, contributed more to the overall light scattering in the mimics than in the WT (14-57% compared to 7% for WT at the same concentration). In N143D and TM, peak 3 with a  $R_h$  between 200-400 nm was detected contributing 6-13% of the total light scattering.

The estimated  $M_w$ -R determined from unfractionated samples were greater for deamidated  $\gamma$ S, particularly for N14D and N76D, than for the WT (Table 1). Since, the  $M_w$ -R is a weight average, even a small amount of a large species or nonglobular protein would increase the  $M_w$ -R more than the predicted  $M_w$  of  $20.9 \times 10^3$  g/mol. The wider distribution of  $R_h$ , greater %Pd and estimated  $M_w$ -R of the deamidated mimics N14D and N76D suggested a small amount of a species larger than a monomer was present, even in Peak 1. All of the mimics had increased light scattering from peak 2 than WT, with N143D and the TM having an even larger species found in peak 3. However, the percent mass of these species predicted from reference globular proteins of similar  $M_w$  suggested their mass was less than 0.5%. Thus, a small number of large light scattering aggregates were detected in several DLS peaks in the deamidation mimics that were not present in WT  $\gamma$ S.

***Surface deamidations increased  $M_w$  of human  $\gamma$ S under native conditions.*** The formation of larger light scattering species by the deamidation mimics was confirmed by SEC-MALS on comparison of WT and TM  $\gamma$ S (Fig. 3A). Both proteins eluted primarily as a monomer at 14.5 mL with a  $M_w$  of  $20.5 \times 10^3$  ( $\pm 0.2\%$ ) g/mol, comparable to the expected  $M_w$  of  $20.9 \times 10^3$  g/mol and with a polydispersity index of 1.000. In addition to the monomer peak, WT and TM  $\gamma$ S exhibited an early eluting peak at 7.5 mL, overlapping with the void peak at 7 mL. This peak of TM was much higher in intensity and broader than that of WT. Similar results were obtained for other deamidation mimics (Fig. S1).

The monomer peaks of WT and TM  $\gamma$ S shown in Fig. 3A were collected and their homogeneity verified by SDS-PAGE. The monomer peaks of each protein were then incubated at 37 °C for 6 days before re-examination by SEC-MALS (Fig. 3B). Incubating TM and, to a lesser degree, WT, generated a minor light scattering peak at 13.5 mL with a  $M_w$  of  $46.8 \times 10^3 (\pm 1.2\%)$  g/mol, corresponding to the approximate  $M_w$  of a  $\gamma$ S dimer (Fig 3B). Moreover, the relative light scattering peak between 6 and 8 mL increased for both TM and WT, while the monomer peak decreased, suggesting the early eluting peak was a large oligomer of the monomer. This peak in TM has greater light intensity than the similarly treated WT  $\gamma$ S.

The high  $M_w$  peak of TM had a  $M_w$  estimated to range from  $3\text{-}9 \times 10^6$  g/mol and contained less than 1% (0.8%) of the total protein. Even less protein eluted in this peak in WT preventing its accurate  $M_w$  determination. Both peaks were confirmed by mass spectrometry to contain  $\gamma$ S. Hence, under native conditions the triple deamidation mimic more heavily favored a small population of conformer that was highly aggregation-prone.

**Effects of surface deamidations on the conformational stability of  $\gamma$ S.** Equilibrium unfolding experiments were performed to measure the thermodynamic stability of deamidated  $\gamma$ S relative to WT  $\gamma$ S (Fig. 4). There are two tryptophan (Trp) residues in the N-terminal domain (N-td) and two in the C-terminal domain (C-td) of  $\gamma$ S. The location of these Trp residues permits their use as probes of the global structure (32). The Trps were selectively excited and their fluorescence measured in the presence of increasing concentrations of the denaturant, GuHCl. The midpoint of unfolding ( $C_m$ ) for WT was 2.58 M, comparable to the previously reported value of 2.3 M(33). The unfolding transition of WT was fitted with a two-state model and a Gibbs free energy of unfolding,  $\Delta G^\circ_{N-U}$ , of 7.58 kcal/mol was derived. The  $C_m$  of the deamidated  $\gamma$ S variants deviated only marginally from that of WT  $\gamma$ S (Fig. 4). However, minor shifts in the early stages of the unfolding transition were observed for N76D and TM  $\gamma$ S (Fig. 4), resulting in changes in  $\Delta G^\circ_{N-U}$  of -1.13 to -2.94 kcal/mol from that of WT. The trend towards a less stable protein at lower denaturant concentrations may reflect unfolding of the less stable N-td (33). The comparable  $C_m$  value with decreased  $\Delta G^\circ_{N-U}$  for N76D and TM indicates a decrease in the  $m$ -value (i.e. the dependence of  $\Delta G^\circ_{N-U}$  on denaturant concentration) which correlates strongly to changes in the amount of solvent accessible area exposed upon unfolding (34). Hysteresis was also observed upon refolding for all deamidation mimics and for WT  $\gamma$ S (Fig. S2). The presence of hysteresis further suggested the tendency for the partially unfolded intermediates to aggregate with differing refolding trends between the proteins.

**During chemical denaturation, partially unfolded intermediates of deamidated  $\gamma$ S formed more rapidly than of WT  $\gamma$ S.** Proteins were unfolded in 2.75 M GuHCl, near the concentration of half equilibrium unfolding ( $C_m$ ) determined in Fig. 4. Following different times of GuHCl exposure, they were subjected to pulse labeling in H/D exchange D<sub>2</sub>O buffer to measure the number of solvent exposed amides (Fig. 5). Three distinct major mass species were detected in WT following exposure to GuHCl for 16 min (Fig. 5A). This region of the spectrum from  $m/z = 906\text{-}918$  contains the +23 charge state of the protein showing mass increases of ~43 Da (Peak 1), ~85 Da (Peak 2), and ~130 Da (Peak 3) in the GuHCl-treated protein, compared to the  $m/z$  of the unlabeled protein at 908.6 (arrows) (spectrum not shown). Peak 1, the major species observed immediately after addition of GuHCl (0 time), results from labeling the most solvent exposed amides in the native  $\gamma$ S structure. In the spectra of TM (Fig. 5B), after 16 min of GuHCl

exposure, Peak 1 was completely lost, and only Peaks 2 and 3 were observed. This indicated that TM unfolds more rapidly than WT, forming the partially unfolded intermediate in Peak 2, and the nearly fully unfolded form in Peak 3 at faster rates. Since the exchange was performed in 90% D<sub>2</sub>O, and back exchange was estimated to be approximately 10%, Peaks 1, 2, and 3, corresponded to the exchange of approximately 31, 62, and 94% of the protein's backbone amides, respectively. The small peak to the right of the arrows, indicating the position of the +23 ion of the unlabeled proteins, may result from aggregated protein that is largely unavailable for H/D exchange.

The percent of Peak 1 (native structure), Peak 2 (partially unfolded intermediate), and peak 3 (nearly fully unfolded protein) was calculated and the time course of their appearance and decrease after 0-32 min of exposure to 2.75 M GuHCl are shown in Fig. 6. At 4, 8, and 16 min, there was a significant decrease in Peak 1 (native structure) in N76D and TM proteins compared to the WT protein (Fig. 6A). The more rapid loss of native structure (Peak 1) in N76D and TM was accompanied by the significantly greater abundance of partially unfolded Peak 2 at 2, 4, and 8 min compared to WT protein (Fig. 6C), and a significantly lower abundance of Peak 2 at 32 min due to the faster accumulation of Peak 3 at 16 and 32 min (Fig. 6E). In comparison, N14D and N143D proteins unfolded at similar rates as the WT protein following GuHCl addition (Fig. 6B, D, and F). While, the identities of the exchanged amide residues in these three peaks were not determined, these results suggested that deamidation in N76D and TM  $\gamma$ S led to more rapid production of partially and fully unfolded forms when  $\gamma$ S was exposed to GuHCl. This observation was also consistent with their greater susceptibility to unfolding in GuHCl under equilibrium conditions (Fig. 4).

***Surface deamidations increased chemically-induced aggregation of  $\gamma$ S.*** To investigate whether deamidation of  $\gamma$ S led to the formation of aggregation-prone, partially unfolded protein states, aggregation of WT and deamidated  $\gamma$ S was monitored under conditions that facilitated unfolding. Proteins were incubated at 37 °C in the presence of 2 M GuHCl, i.e. a denaturant concentration corresponding to the beginning of the unfolding transition determined in Fig. 4. For WT there was a sigmoidal increase in solution turbidity and ThT binding over time which together indicated that the protein was forming aggregates with a degree of ordered fibrillar structure, as reported previously for  $\gamma$ D-crystallin under comparable conditions (35). For N14D and N143D there was a minor but significant ( $p < 0.05$ ) increase in aggregation propensity relative to WT, as shown by higher rates of aggregation (Fig. 7, inset) as well as higher levels of turbidity and ThT binding at the end of the incubation period (Fig. 7). In contrast, N76D showed a considerably greater aggregation propensity, consistent with its reduced stability (Fig. 4) and increased solvent accessibility under denaturing conditions (Fig. 6). The effects of deamidation at each site were accumulative with TM exhibiting the highest propensity to aggregate on the basis of aggregation rate. Despite the higher aggregation rate and final ThT binding, the solution turbidity of TM was disproportionally low, indicating a shift to more elongated (i.e. rod-like) structures (36). It is concluded that deamidation of  $\gamma$ S, particularly at N76, promotes partial unfolding which thereby leads to enhanced aggregation.

***Surface deamidations increased thermally-induced aggregation of  $\gamma$ S.*** The enhanced aggregation propensity of deamidated  $\gamma$ S was further investigated under heat stress at 70°C (Fig. 8)(26). The N14D, N76D and TM mimics exhibited significantly greater turbidity after 5



minutes of exposure to elevated temperature when compared to WT and N143D  $\gamma$ S (Fig. 8). Additionally, the rates of aggregation of N14D, N76D and TM were significantly greater than WT and N143D. The order of temperature stability from most to least stable was WT/N143D > N14D/N76D > TM. Deamidation at N14 and N76 located on the surface of the N-td affected the thermal-stability more than the deamidation at N143 on the surface of the C-td and the effects were cumulative, i.e. in TM.

WT and TM  $\gamma$ S were next incubated at 65 °C with and without the presence of a 2:1 molar subunit ratio of the molecular chaperone  $\alpha$ A-crystallin ( $\alpha$ A) to probe differences introduced from surface deamidations (Fig. 9). Both WT alone and WT in the presence of  $\alpha$ A exhibited little change in turbidity with time as measured by light scattering at 405 nm over the course of the experiment. As expected, TM alone increased in turbidity over 15 minutes of heating, after which there was no further change. The lower turbidity value compared to Fig. 8 is due to the lower temperature that was used in that experiment. When incubated in the presence of  $\alpha$ A, no increase in turbidity of TM occurred during the first 15 minutes, before aggregation commenced. Samples were analyzed by SDS-PAGE and mass spectrometry and it was determined that the precipitated proteins resulting from heating of TM alone and TM with  $\alpha$ A, contained  $\gamma$ S and  $\gamma$ S plus  $\alpha$ A, respectively (Fig. S4). The large error bar of the TM +  $\alpha$ A sample at 60 min reflects the variability from the sample precipitating and large aggregates that formed. These results suggest that  $\alpha$ A forms an unstable complex with deamidated  $\gamma$ S that ultimately leads to rapid aggregation and co-precipitation.

## Discussion

While cataracts are the major cause of blindness worldwide, the molecular mechanisms that lead to the formation of large light scattering aggregates that cause lens opacities are not known. The major proteins in the human eye lens, crystallins, are extensively modified during aging and cataract formation. Yet, a direct link between the accumulation of age-related modified crystallins and cataracts has not been well established. We present evidence here that one of the most prevalent modifications, deamidation of labile Asn residues, directly causes protein aggregation and present data that support a mechanism via the aggregation of partially unfolded intermediates as depicted in Fig. 10.

Deamidations at the labile Asn residues at positions 14, 76, and 143 in human  $\gamma$ S led to large light scattering aggregates with a  $R_h$  greater than 10 nm that contributed 14-60% of the total scattering intensity compared to 7% for WT under the same conditions. Similar methods were used to detect light scattering in the P23T mutation in human  $\gamma$ D-crystallin (similar to the P23V congenital cataract mutation), which showed < 5% of the total scattering intensity (37).

Incubating the isolated  $\gamma$ S monomer under physiological conditions led to increased aggregate formation with a corresponding decrease in monomer, supporting that deamidation directly led to the protein's aggregation. It had previously been reported that deamidation at N76 and N143 together resulted in greater attractive interactions and that deamidation at N76 led to a tendency towards dimer formation without perturbing the structure or stability of  $\gamma$ S (7). Taken together, these findings are consistent with deamidation leading to aggregation and thus, contributing to the insoluble deamidated crystallins isolated from cataractous lenses.

In these *in vitro* studies, the levels of deamidation-induced aggregation were < 1% of the total protein concentration, yet contributed to approximately half the light scattering intensity of the total protein depending on the deamidation site, i.e. only less than 1% of deamidated aggregate scattered a significant amount of light. In the crowded cellular milieu of the lens, this level of deamidation-induced aggregation is likely significant.

In order to determine if deamidation-induced aggregation was due to changes in stability, structures were probed during chemical and thermal-denaturation. Only small differences in equilibrium chemical-induced unfolding between WT and deamidated mimics were observed, suggesting that the overall structure of deamidated  $\gamma$ S was comparable to that of WT (7, 26). Similar findings have been reported of none or little differences in the overall structure between several congenital cataract mutants (4, 5, 8, 9, 38-40). An emerging theme of the role of age-related modifications in crystallin structure is that these modifications do not appear to alter the overall structure of crystallins significantly *in vitro*.

However, during chemical kinetic unfolding of  $\gamma$ S, a partially unfolded intermediate and a nearly fully unfolded intermediate were observed by H/D exchange that formed more rapidly in N76D and TM than WT. At a slightly lower GuHCL concentration the formation of these intermediates led to their rapid aggregation, suggesting that the partially unfolded intermediates were the species that were aggregating. The specific backbone amides undergoing rapid exchange in this intermediate have not yet been determined. The partially unfolded intermediate likely corresponds to the unfolding of the first Greek Key motif with approximately a 40 Da increase in mass in the N-td that was more prone to aggregate in the deamidated  $\gamma$ S mimics. Others have reported that the N-td would unfold first followed by the C-td (33, 35, 41). Likewise, the N-td deamidations at N14 and N76 impacted  $\gamma$ S more than the deamidation at N143 in the more stable C-td. Illustrating this is the observation that mutation at, at least, seven positions in  $\gamma$ S is associated with early-onset cataract, whereas no inheritable mutations have so far been identified in the C-td (42).

During thermally-induced unfolding, deamidated crystallins rapidly aggregated supporting the chemical-unfolding findings. The mimic containing all three deamidation sites had the greatest effect on the structure of  $\gamma$ S, with the N143D mimic in the C-td having only a minor effect on thermal-stability. Both deamidated mimics, N14D and N76D, in the N-td had a greater effect on thermal-stability than N143D in the C-td, further reflecting differences between the domains. Deamidated crystallins were also ineffectively prevented from precipitation by the chaperone action of  $\alpha$ -crystallin with both proteins precipitating together (43-45). Similar effects of deamidation on the domain-domain interface in  $\beta$ B2 and  $\gamma$ D have been reported (7, 21, 46). Our data suggest that once a threshold is reached where deamidation accumulates enough to unfold a protein, a cascade of nucleation-dependent aggregation and subsequent precipitation is initiated (47). More exposed hydrophobic regions would be present in deamidated  $\gamma$ S than in WT for interaction with  $\alpha$ A, given the lower temperature at which it aggregated. Several recent reports of deamidation at N76 in  $\gamma$ S support these findings of decreased stability and ineffective chaperoning by  $\alpha$ -crystallin due to the relative rapid aggregation of deamidated  $\gamma$ S (43, 45, 48-50).

In earlier studies we detected changes in conformational dynamics in deamidated  $\gamma$ S, which may help explain the results observed in the present study. These changes in dynamics were detected by differences in H/D exchange by NMR spectroscopy throughout the protein, i.e. at sites distant from the mutations including the  $\beta$ -helical regions in both the N- and C-tds (27). Noted changes were in  $\beta$ -strands in the first Greek Key motif near deamidation sites at N14 and N76. Increased exchange suggested increased solvent accessibility with a time frame of minutes to days of motion. More of the buried residues in these protein regions became exposed in the deamidated mimics during this time than in WT  $\gamma$ S. The greater exposure of core residues and accessibility of other residues over time can explain the increased tendency to aggregate during both thermal and chemical unfolding of the deamidated mimics, particularly of the triple mimic.

$\gamma$ -crystallins are stable, globular proteins that have considerable differences in Gibbs free energies between their native and unfolded states, mainly due to the energetically favorable burial of hydrophobic amino acid residues upon folding (33, 41). Complementary to this, the surface of  $\gamma$ -crystallins is dominated by hydrophilic and charged residues, particularly in loop regions, that form favorable hydrogen bonds with water, thereby providing solubility and stability (19). Human  $\gamma$ S has ten Asp and fourteen Glu residues; two-thirds of these acidic residues have side-chain solvent exposures ranging between 40 to 80% with an average of ~55%, since solvation of charged residues is energetically favorable. Comparably, N14, N76 and N143 have solvent exposures of approximately 55, 58 and 60%, respectively. Ionized, aspartyl residues resulting from deamidation form stronger bonds with water than their uncharged, asparagine precursors. However, because surface residues are solvent-accessible in both native and unfolded forms of the protein, alterations in surface charge generally have a minor effect on net protein stability (51).

We attribute the increased aggregation propensity of the deamidation mimics to conformational fluctuations that expose otherwise slow-exchanging core residues to water, triggering unfolding and intermolecular association (27, 52). Additionally, and distinct from other sites of modification (e.g. N14, N143), deamidation at N76 may result in a non-native, surface salt bridge between the newly introduced anionic carboxylate and the cationic ammonium of R78. The minor increase in enthalpy accompanying salt bridge formation may come at a cost of conformational entropy as reflected by the conformational flexibility in the higher  $S^2$  values in the loop region intervening  $\alpha$ -helix a2 and  $\beta$ -strand c2 and the reduced thermodynamic stability of N76D (Fig. 8) (27, 53, 54).

In summary, we have mimicked the aging process in crystallins by introducing deamidations via mutation. ‘Aged’ crystallins tend to partially unfold, exposing hydrophobic residues, as they accumulate deamidations. Exposed hydrophobicity leads to crystallin aggregation and the formation of large light scattering bodies that ultimately precipitate, causing the opacification of the eye lens associated with cataract (Fig. 10).

### **Acknowledgements:**

The authors thank Wyatt Technology Inc. for use of the Dynapro Plate Reader.

**Funding Information:** This research was supported in part by an Australian National Health and Medical Research Council Project Grant to J.A.C. (#1068087), by the National Institutes of Health Grants EY027012 to KJL and EY027768 to KJL and LLD. Mass spectrometry support includes Ophthalmology Proteomics Core NIH P30 EY010572 and S10 OD012246 to Oregon Health and Science University.

**Author Contribution Statement:**

CJV, DCT, CM, and KH performed experiments. LLD designed and performed mass spectrometry experiments and wrote methods. CJV, DCT and JAC designed experiments and wrote manuscript. KJL designed overall experiments and wrote manuscript.

**References:**

1. Delaye M, Tardieu A. Short-range order of crystallin proteins accounts for eye lens transparency. *Nature*. 1983;302(5907):415-7.
2. Truscott RJ, Friedrich MG. The etiology of human age-related cataract. Proteins don't last forever. *Biochim Biophys Acta*. 2016;1860(1 Pt B):192-8.
3. Congdon N, Vingerling JR, Klein BE, West S, Friedman DS, Kempen J, et al. Prevalence of cataract and pseudophakia/aphakia among adults in the United States. *Arch Ophthalmol*. 2004;122(4):487-94.
4. Ji F, Jung J, Gronenborn AM. Structural and biochemical characterization of the childhood cataract-associated R76S mutant of human  $\gamma$ D-crystallin. *Biochemistry*. 2012;51(12):2588-96.
5. Ji F, Jung J, Koharudin LM, Gronenborn AM. The human W42R gammaD-crystallin mutant structure provides a link between congenital and age-related cataracts. *J Biol Chem*. 2013;288(1):99-109.
6. Wilmarth PA, Tanner S, Dasari S, Nagalla SR, Riviere MA, Bafna V, et al. Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility? *J Proteome Res*. 2006;5(10):2554-66.
7. Pande A, Mokhor N, Pande J. Deamidation of Human gammaS-Crystallin Increases Attractive Protein Interactions: Implications for Cataract. *Biochemistry*. 2015;54(31):4890-9.
8. Boatz JC, Whitley MJ, Li M, Gronenborn AM, van der Wel PCA. Cataract-associated P23T  $\gamma$ D-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH. *Nat Commun*. 2017;8:15137-.
9. Ji F, Koharudin LM, Jung J, Gronenborn AM. Crystal structure of the cataract-causing P23T gammaD-crystallin mutant. *Proteins*. 2013;81(9):1493-8.
10. Treweek TM, Meehan S, Ecroyd H, Carver JA. Small heat-shock proteins: important players in regulating cellular proteostasis. *Cell Mol Life Sci*. 2015;72(3):429-51.
11. Carver JA, Ecroyd H, Truscott RJW, Thorn DC, Holt C. Proteostasis and the Regulation of Intra- and Extracellular Protein Aggregation by ATP-Independent Molecular Chaperones: Lens  $\alpha$ -Crystallins and Milk Caseins. *Acc Chem Res*. 2018;51(3):745-52.
12. Basak A, Bateman O, Slingsby C, Pande A, Asherie N, Ogun O, et al. High-resolution X-ray crystal structures of human gammaD crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. *J Mol Biol*. 2003;328(5):1137-47.
13. Bax B, Lapatto R, Nalini V, Driessen H, Lindley PF, Mahadevan D, et al. X-ray analysis of beta B2-crystallin and evolution of oligomeric lens proteins. *Nature*. 1990;347(6295):776-80.

14. Van Montfort RL, Bateman OA, Lubsen NH, Slingsby C. Crystal structure of truncated human betaB1-crystallin. *Protein Sci.* 2003;12(11):2606-12.
15. Lampi KJ, Murray MR, Peterson MP, Eng BS, Yue E, Clark AR, et al. Differences in solution dynamics between lens beta-crystallin homodimers and heterodimers probed by hydrogen-deuterium exchange and deamidation. *Biochim Biophys Acta.* 2016;1860(1 Pt B):304-14.
16. Lampi KJ, Ma Z, Hanson SR, Azuma M, Shih M, Shearer TR, et al. Age-related changes in human lens crystallins identified by two-dimensional electrophoresis and mass spectrometry. *Exp Eye Res.* 1998;67(1):31-43.
17. Dasari S, Wilmarth PA, Reddy AP, Robertson LJ, Nagalla SR, David LL. Quantification of isotopically overlapping deamidated and 18O-labeled peptides using isotopic envelope mixture modeling. *J Proteome Res.* 2009;8(3):1263-70.
18. Takemoto L, Boyle D. Increased deamidation of asparagine during human senile cataractogenesis. *Mol Vis.* 2000;6:164-8.
19. Lapko VN, Purkiss AG, Smith DL, Smith JB. Deamidation in human gamma S-crystallin from cataractous lenses is influenced by surface exposure. *Biochemistry.* 2002;41(27):8638-48.
20. Fujii N, Takata T, Fujii N, Aki K. Isomerization of aspartyl residues in crystallins and its influence upon cataract. *Biochim Biophys Acta.* 2016;1860(1 Pt B):183-91.
21. Hooi MY, Raftery MJ, Truscott RJ. Age-dependent deamidation of glutamine residues in human gammaS crystallin: deamidation and unstructured regions. *Protein Sci.* 2012;21(7):1074-9.
22. Hooi MY, Raftery MJ, Truscott RJ. Racemization of two proteins over our lifespan: deamidation of asparagine 76 in gammaS crystallin is greater in cataract than in normal lenses across the age range. *Invest Ophthalmol Vis Sci.* 2012;53(7):3554-61.
23. Kingsley CN, Brubaker WD, Markovic S, Diehl A, Brindley AJ, Oschkinat H, et al. Preferential and specific binding of human alphaB-crystallin to a cataract-related variant of gammaS-crystallin. *Structure.* 2013;21(12):2221-7.
24. Hanson SR, Hasan A, Smith DL, Smith JB. The major in vivo modifications of the human water-insoluble lens crystallins are disulfide bonds, deamidation, methionine oxidation and backbone cleavage. *Exp Eye Res.* 2000;71(2):195-207.
25. Hains PG, Truscott RJ. Post-translational modifications in the nuclear region of young, aged, and cataract human lenses. *J Proteome Res.* 2007;6(10):3935-43.
26. Ray NJ, Hall D, Carver JA. Deamidation of N76 in human gammaS-crystallin promotes dimer formation. *Biochim Biophys Acta.* 2016;1860(1 Pt B):315-24.
27. Forsythe HM, Vetter CJ, Jara KA, Reardon PN, David LL, Barbar EJ, et al. Altered Protein Dynamics and Increased Aggregation of Human  $\gamma$ S-Crystallin Due to Cataract-Associated Deamidations. *Biochemistry.* 2019;58(40):4112-24.
28. Takata T, Oxford JT, Brandon TR, Lampi KJ. Deamidation alters the structure and decreases the stability of human lens betaA3-crystallin. *Biochemistry.* 2007;46(30):8861-71.
29. Cox D, Selig E, Griffin MDW, Carver JA, Ecroyd H. Small Heat-shock Proteins Prevent  $\alpha$ -Synuclein Aggregation via Transient Interactions and Their Efficacy Is Affected by the Rate of Aggregation. *J Biol Chem.* 2016;291(43):22618-29.
30. Thorn DC, Grosas AB, Mabbitt PD, Ray NJ, Jackson CJ, Carver JA. The Structure and Stability of the Disulfide-Linked  $\gamma$ S-Crystallin Dimer Provide Insight into Oxidation Products Associated with Lens Cataract Formation. *J Mol Biol.* 2019;431(3):483-97.



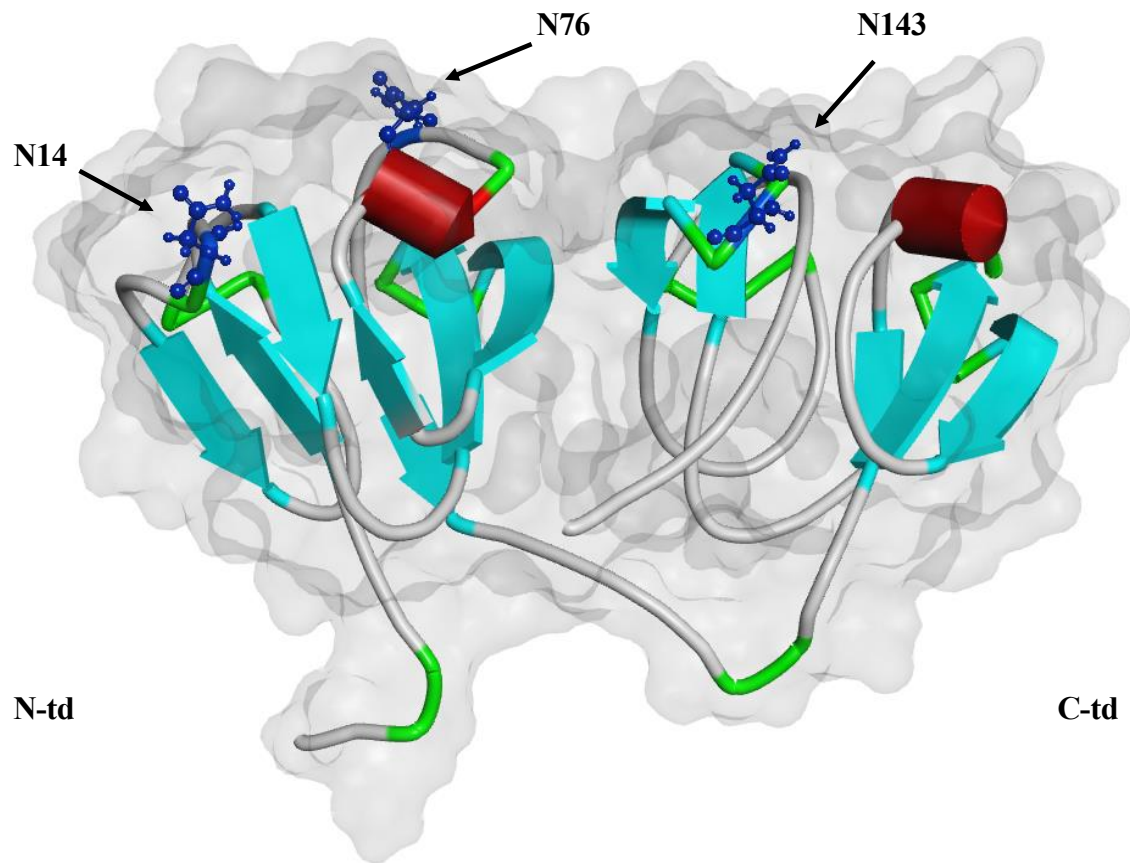
31. Jiang J, Golchert KJ, Kingsley CN, Brubaker WD, Martin RW, Mukamel S. Exploring the aggregation propensity of  $\gamma$ S-crystallin protein variants using two-dimensional spectroscopic tools. *J Phys Chem B*. 2013;117(46):14294-301.
32. Serebryany E, Takata T, Erickson E, Schafheimer N, Wang Y, King JA. Aggregation of Trp > Glu point mutants of human gamma-D crystallin provides a model for hereditary or UV-induced cataract. *Protein Sci*. 2016;25(6):1115-28.
33. Mills IA, Flaugh SL, Kosinski-Collins MS, King JA. Folding and stability of the isolated Greek key domains of the long-lived human lens proteins gammaD-crystallin and gammaS-crystallin. *Protein Sci*. 2007;16(11):2427-44.
34. Myers JK, Pace CN, Scholtz JM. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci*. 1995;4(10):2138-48.
35. Kosinski-Collins MS, King J. In vitro unfolding, refolding, and polymerization of human gammaD crystallin, a protein involved in cataract formation. *Protein Sci*. 2003;12(3):480-90.
36. Hall D, Zhao R, Dehlsen I, Bloomfield N, Williams SR, Arisaka F, et al. Protein aggregate turbidity: Simulation of turbidity profiles for mixed-aggregation reactions. *Anal Biochem*. 2016;498:78-94.
37. Evans P, Wyatt K, Wistow GJ, Bateman OA, Wallace BA, Slingsby C. The P23T cataract mutation causes loss of solubility of folded gammaD-crystallin. *J Mol Biol*. 2004;343(2):435-44.
38. Jung J, Byeon I-JL, Wang Y, King J, Gronenborn AM. The structure of the cataract-causing P23T mutant of human gammaD-crystallin exhibits distinctive local conformational and dynamic changes. *Biochemistry*. 2009;48(12):2597-609.
39. Khago D, Wong EK, Kingsley CN, Freitas JA, Tobias DJ, Martin RW. Increased hydrophobic surface exposure in the cataract-related G18V variant of human  $\gamma$ S-crystallin. *Biochim Biophys Acta*. 2016;1860(1 Pt B):325-32.
40. Kingsley CN, Brubaker WD, Markovic S, Diehl A, Brindley AJ, Oschkinat H, et al. Preferential and specific binding of human  $\alpha$ B-crystallin to a cataract-related variant of  $\gamma$ S-crystallin. *Structure*. 2013;21(12):2221-7.
41. Mills-Henry IA, Thol SL, Kosinski-Collins MS, Serebryany E, King JA. Kinetic Stability of Long-Lived Human Lens gamma-Crystallins and Their Isolated Double Greek Key Domains. *Biophys J*. 2019;117(2):269-80.
42. Zhang K, Zhu X, Lu Y. The Proteome of Cataract Markers: Focus on Crystallins. *Adv Clin Chem*. 2018;86:179-210.
43. Lampi KJ, Fox CB, David LL. Changes in solvent accessibility of wild-type and deamidated betaB2-crystallin following complex formation with alphaA-crystallin. *Exp Eye Res*. 2012;104:48-58.
44. Michiel M, Duprat E, Skouri-Panet F, Lampi JA, Tardieu A, Lampi KJ, et al. Aggregation of deamidated human betaB2-crystallin and incomplete rescue by alpha-crystallin chaperone. *Exp Eye Res*. 2010;90(6):688-98.
45. Lampi KJ, Kim YH, Bachinger HP, Boswell BA, Lindner RA, Carver JA, et al. Decreased heat stability and increased chaperone requirement of modified human betaB1-crystallins. *Mol Vis*. 2002;8:359-66.
46. Flaugh SL, Mills IA, King J. Glutamine deamidation destabilizes human gammaD-crystallin and lowers the kinetic barrier to unfolding. *J Biol Chem*. 2006;281(41):30782-93.

47. Lampi KJ, Wilmarth PA, Murray MR, David LL. Lens beta-crystallins: the role of deamidation and related modifications in aging and cataract. *Prog Biophys Mol Biol*. 2014;115(1):21-31.
48. Ray NJ, Hall D, Carver JA. Deamidation of N76 in human gammaS-crystallin promotes dimer formation. *Biochim Biophys Acta*. 2015.
49. Carver JA, Lindner RA, Lyon C, Canet D, Hernandez H, Dobson CM, et al. The interaction of the molecular chaperone alpha-crystallin with unfolding alpha-lactalbumin: a structural and kinetic spectroscopic study. *J Mol Biol*. 2002;318(3):815-27.
50. Treweek TM, Rekas A, Lindner RA, Walker MJ, Aquilina JA, Robinson CV, et al. R120G alphaB-crystallin promotes the unfolding of reduced alpha-lactalbumin and is inherently unstable. *FEBS J*. 2005;272(3):711-24.
51. Sanchez-Ruiz JM, Makhatadze GI. To charge or not to charge? *Trends Biotechnol*. 2001;19(4):132-5.
52. Wintrode PL, Zhang D, Vaidehi N, Arnold FH, Goddard WA, 3rd. Protein dynamics in a family of laboratory evolved thermophilic enzymes. *J Mol Biol*. 2003;327(3):745-57.
53. Takano K, Tsuchimori K, Yamagata Y, Yutani K. Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry*. 2000;39(40):12375-81.
54. Strop P, Mayo SL. Contribution of surface salt bridges to protein stability. *Biochemistry*. 2000;39(6):1251-5.

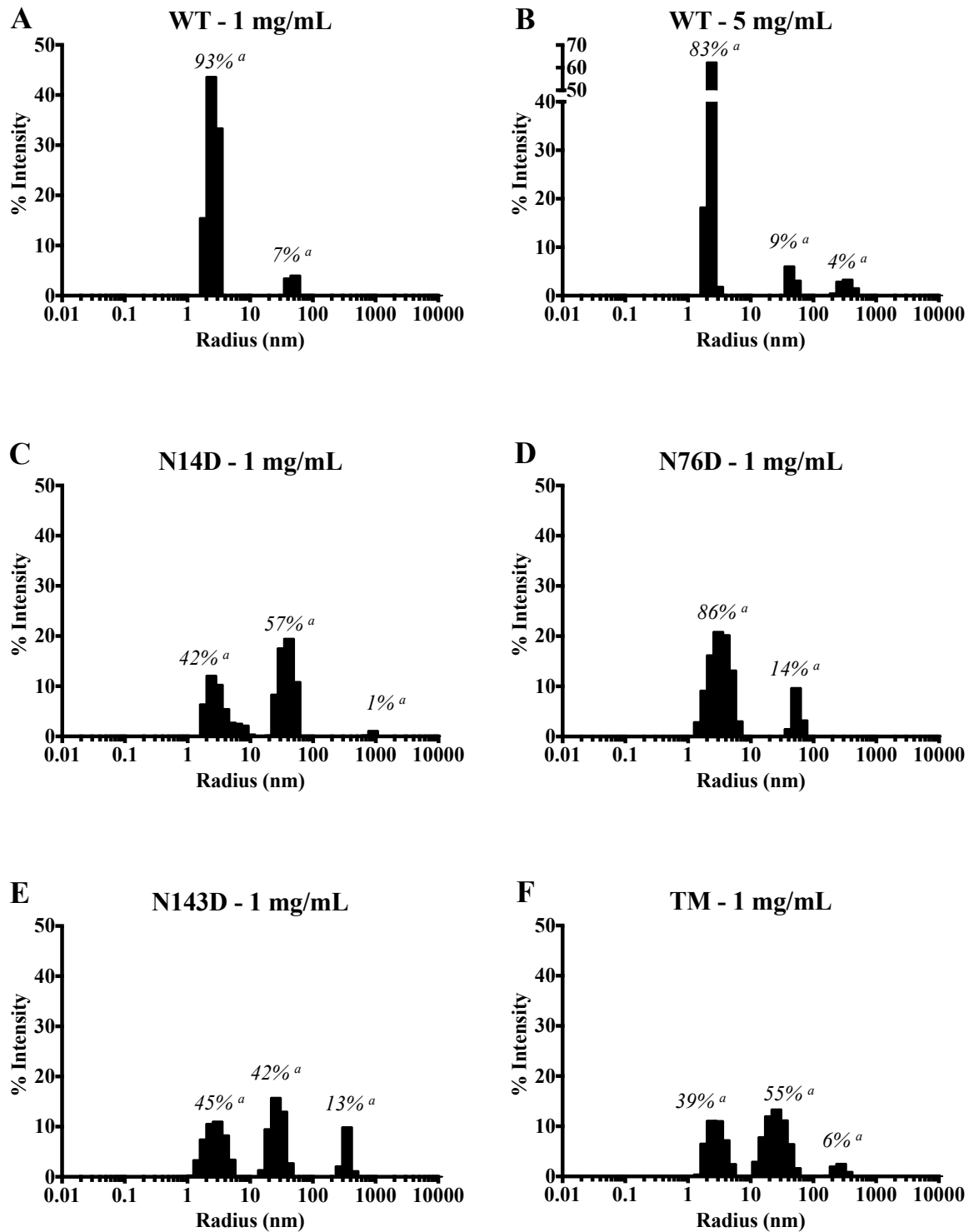
**Table 1. Sizes of  $\gamma$ S in Peak 1 determined by dynamic light scattering**

$\gamma$ S	Concentration	Radius (nm)	% Pd	Mw-R
	mg/mL	Mean (SEM)	Mean (SEM)	Mean (SEM)
<b>WT</b>	1	2.5 (0.2)	19 (4)	28.9 (3.26)
	5	2.4 (0.2)	6* (2)	25.0 (0.8)
<b>N14D</b>	1	3.2 (0.5)	37* (6)	49.2** (5.6)
<b>N76D</b>	1	3.3 (0.3)	35* (3)	55.0** (3.8)
<b>N143D</b>	1	2.7 (0.2)	29 (2)	32.7 (3.6)
<b>TM</b>	1	2.9 (0.3)	22 (4)	36.6 (2.5)

The radius, % polydispersity (%Pd), and estimated weight average molar mass (Mw-R) of  $\gamma$ S in peak 1 (1-10 nm) from Fig. 2. Significant differences in Mw-R between WT (1 mg/mL) and deamidated  $\gamma$ S were determined using a T-test are denoted with  $p \leq 0.05$  (\*) and  $p \leq 0.01$  (\*\*), (n = 3-5).



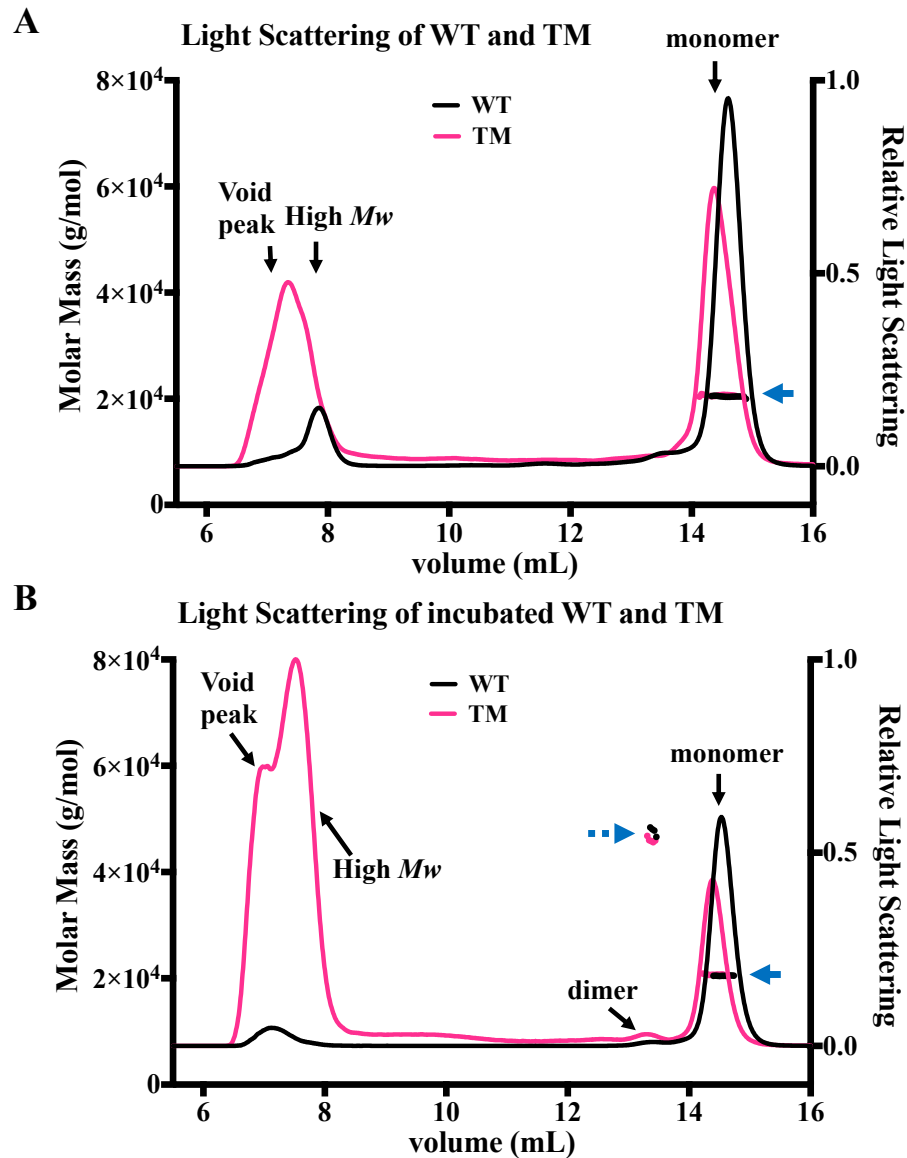
**Figure 1: Location of Asn residues 14, 76, and 143 in human  $\gamma$ S.** In order to mimic deamidation, site-directed mutagenesis was used to replace Asn with Asp at residues 14, 76, and 143. Asparagine residues 14 and 76 are located on surface of the N-terminal domain (N-td) and residue 143 is located on the surface of the C-terminal domain (C-td). Residues of interest are shown in ball and stick structure. Grey area represents solvent accessibility with a probe radius of 1.4, cyan arrows represent  $\beta$ -sheets, red cylinders represent  $\alpha$ -helices and green represents coil and turn structures as generated with Discovery Studio Visualizer (BIOVIA, San Diego, CA). (PDB: 2M3T)



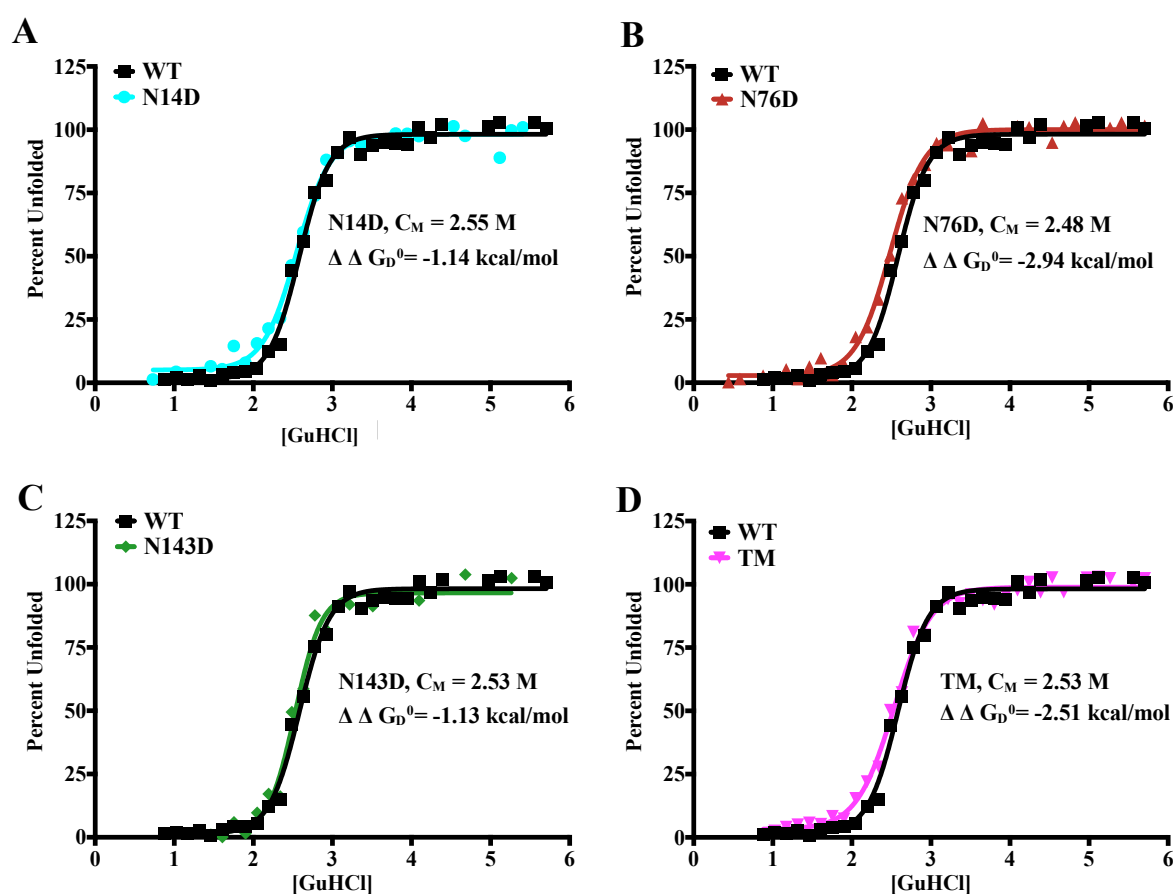
<sup>a</sup> (% , peak DLS/ total DLS in sample)



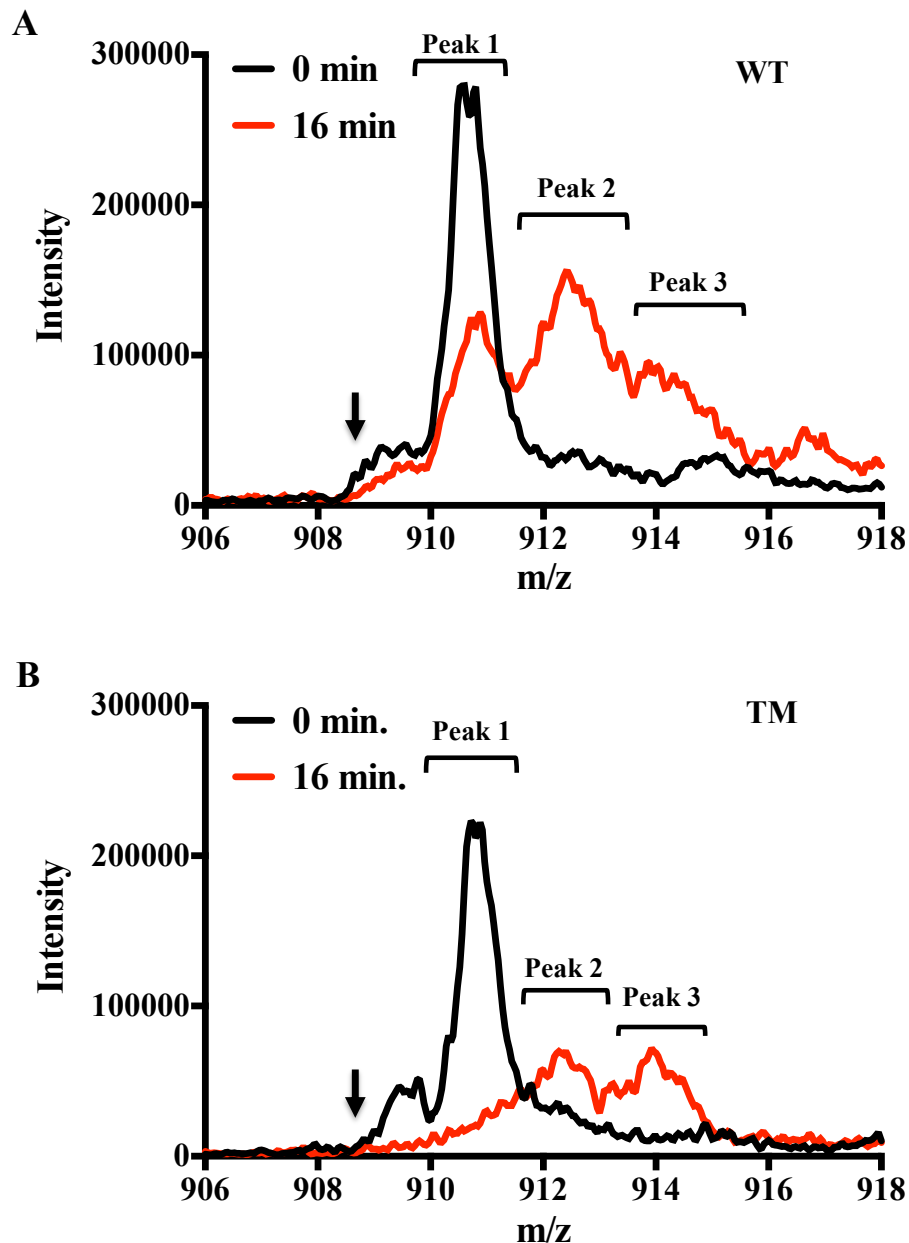
**Figure 2: Effect of deamidation on radii of hydration,  $R_h$ , of  $\gamma$ S.** The percent of dynamic light scattering intensity (DLS) was plotted for the size distributions between 1-10 nm (peak 1), 10-100 nm (peak 2), and 100-1000 nm (peak 3) of WT and deamidated  $\gamma$ S. **A and B.** The % intensity of WT species at 1 mg/mL and 5 mg/mL respectively. **C, D, E and F.** The % intensity of N14D, N76D, N143D and N14D/N76D/N143D (Triple Mutant, TM) at 1 mg/mL. Data collected at 25 °C. <sup>a</sup>The % DLS for each peak relative to the total DLS for that sample. Data are representative of three independent experiments.



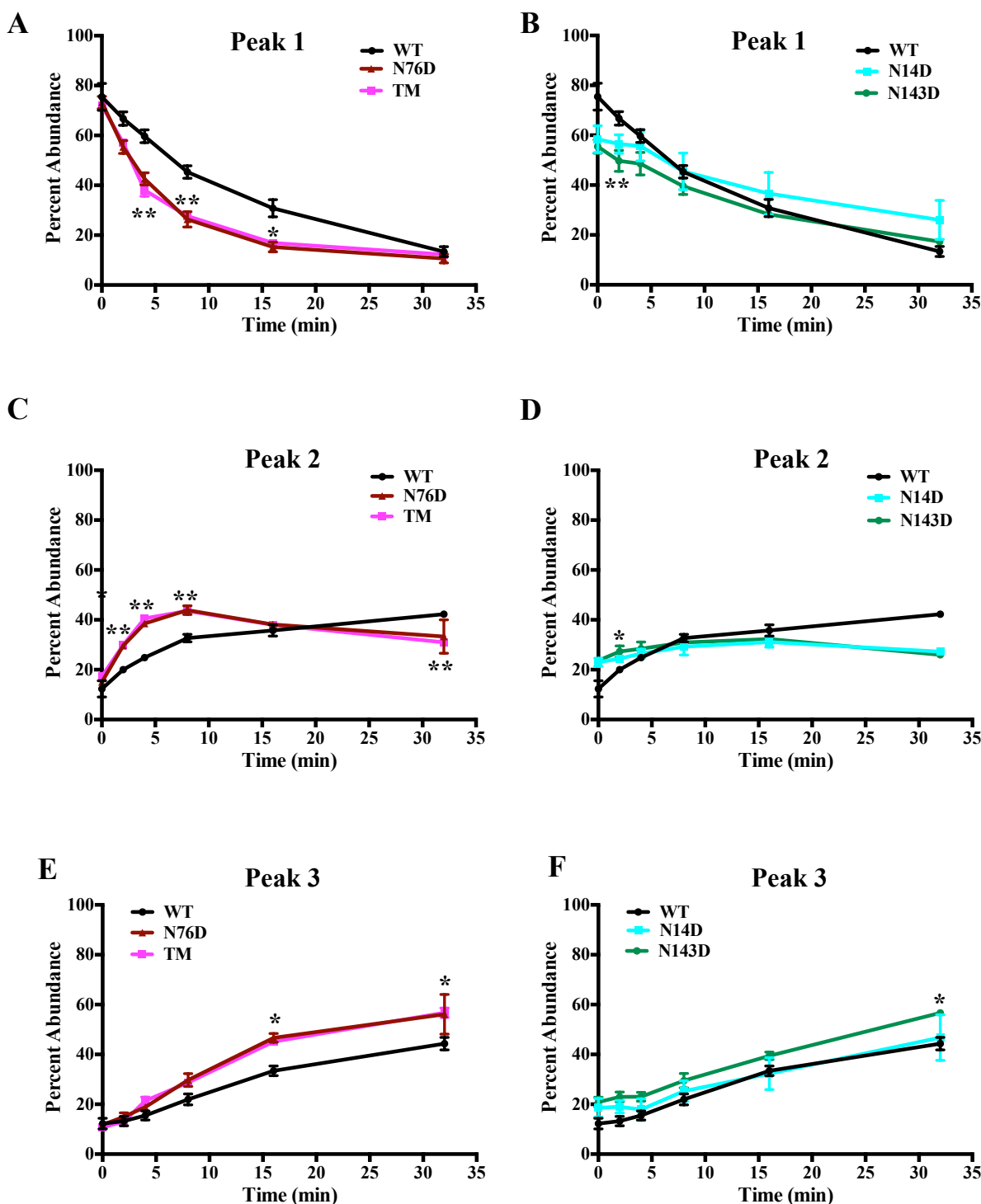
**Figure 3: Relative light scattering (LS) of WT and deamidated  $\gamma$ S.** **A.** Light scattering of WT and TM  $\gamma$ S during size exclusion chromatography in-line with multiangle light scattering (SEC-MALS). The right axis is relative Rayleigh LS of WT (solid, black line) and TM (solid, magenta line). The left axis is the molar mass (noted by blue arrows). Monomer peaks eluted at 14.5 mL and high  $M_w$  peaks eluted at 7.5 mL. Molar masses for the monomer peaks were  $20.5 \times 10^3$  g/mol (solid, blue arrow) compared to the predicted  $M_w$  of  $20.9 \times 10^3$ . **B.** Light scattering of incubated WT and TM  $\gamma$ S monomers. The monomer peaks isolated in panel A were incubated at 37 °C for 6 days and then reanalyzed by SEC-MALS. Monomer peaks were detected with  $20.6 \times 10^3$  g/mol (solid, blue arrow) and dimer peaks eluted at 13.5 mL with  $46.8 \times 10^3$  g/mol (dashed, blue arrow). Molar mass for the high  $M_w$  peak for the TM in panel B was estimated to range from  $3\text{-}9 \times 10^6$  g/mol (data not shown).



**Figure 4: Equilibrium unfolding of WT and deamidated  $\gamma$ S during incubation in GuHCl.** Unfolding of  $\gamma$ S was measured by changes in tryptophan fluorescence following an overnight incubation in increasing concentrations of GuHCl. **A-D.** Unfolding curves for N14D, N76D, N143D and TM, with the WT curve superimposed. The change in Gibbs free energy,  $\Delta \Delta G_D^0$  (kcal/mol), of unfolding was calculated using the difference between the  $\Delta G_D^0$  of WT (8.59 kcal/mol), and the  $\Delta G_D^0$  of each deamidated  $\gamma$ S. The  $C_M$  for WT was 2.58 M.

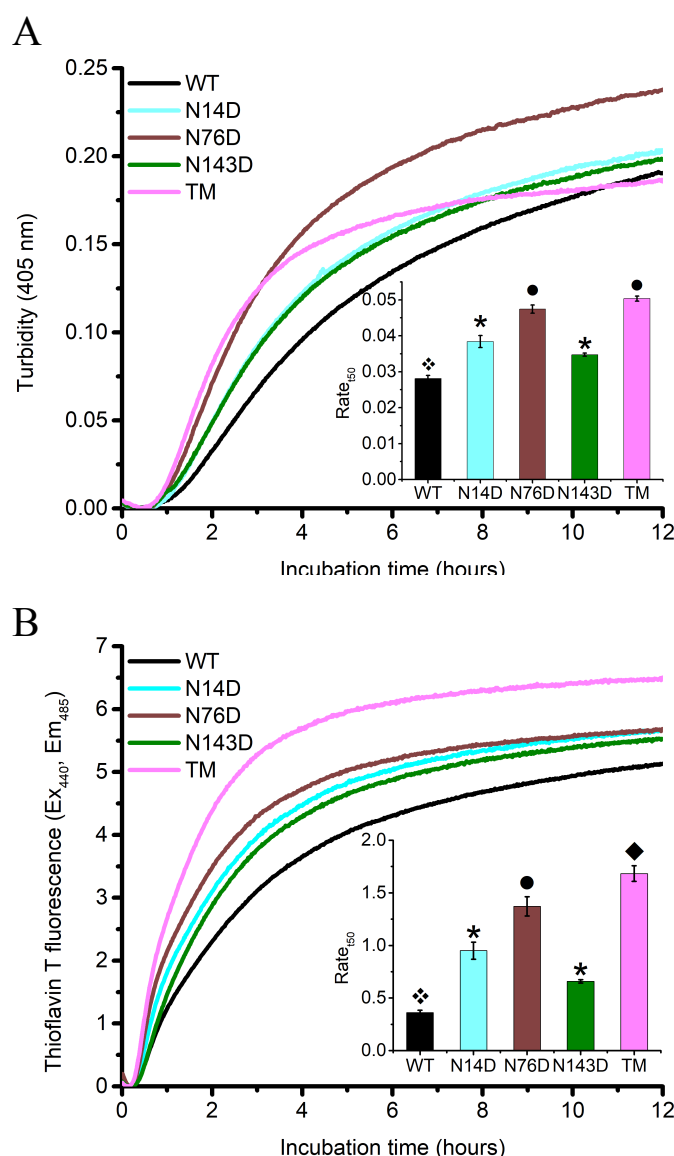


**Figure 5: Deamidation increases the rate of  $\gamma$ S chemical unfolding.** Mass spectrum of the +23 charge state of WT and TM  $\gamma$ S following partial unfolding in 2.75 M GuHCl. Proteins were unfolded in 2.75 M GuHCl for 0 and 16 min and subjected to pulse labeling in hydrogen deuterium (H/D) exchange  $D_2O$  buffer. Three major peaks were detected after 16 min in GuHCl with mass increases of 36-43 Da (peak 1), 83-85 Da (peak 2), and 120-130 Da (peak 3) over the unlabeled proteins (position of arrow, spectra not shown). **A.** The spectrum of WT at 0 min is overlaid with WT at 16 min. **B.** The spectrum of TM at 0 min is overlaid with TM at 16 min. In both panels A and B, there is a decrease in Peak 1 (folded species) and an increase in Peak 2 (partially unfolded intermediate) and Peak 3 (nearly fully unfolded species). Spectra are averaged from 3 independent experiments.

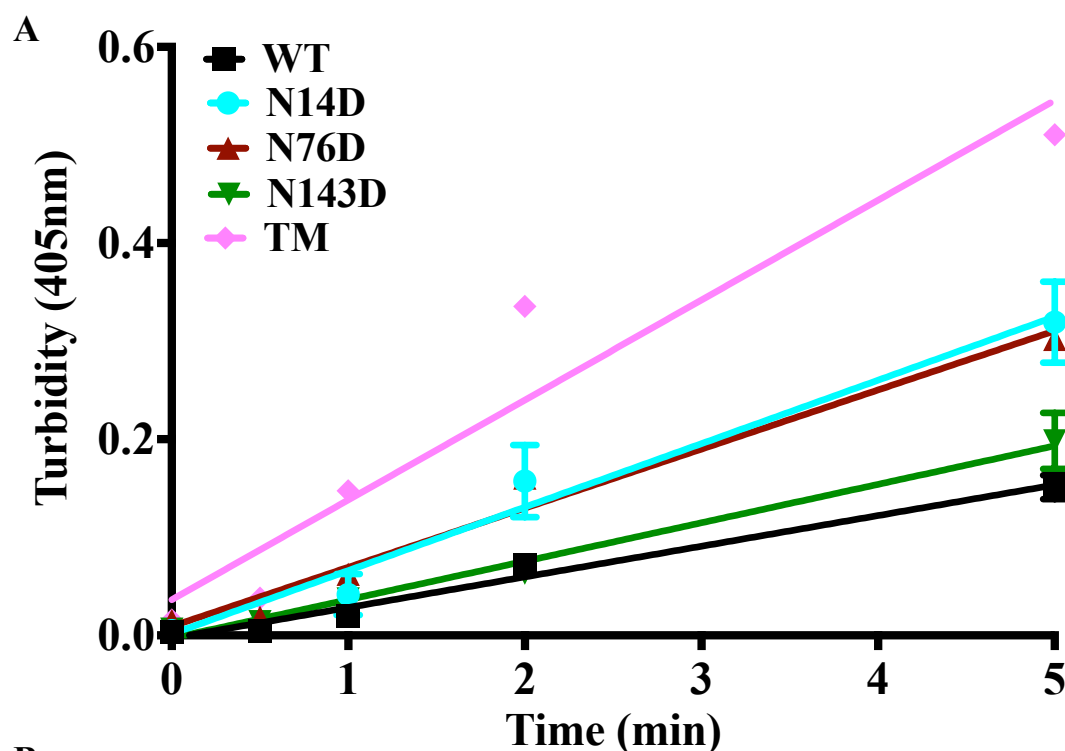


**Figure 6: Changes in relative abundance of WT and deamidated  $\gamma$ S intermediates during chemical unfolding.** Percent area under the peak for each species of  $\gamma$ S protein as shown in Fig. 5, calculated using raw spectra from pulse labeled protein in  $D_2O$  containing H/D exchange buffer. All data presented as percent abundance relative to total area of all peaks. Significant differences from WT with p-value  $\leq 0.01$  are denoted by \*\* and p-value  $\leq 0.05$  denoted by \*, (n = 3).





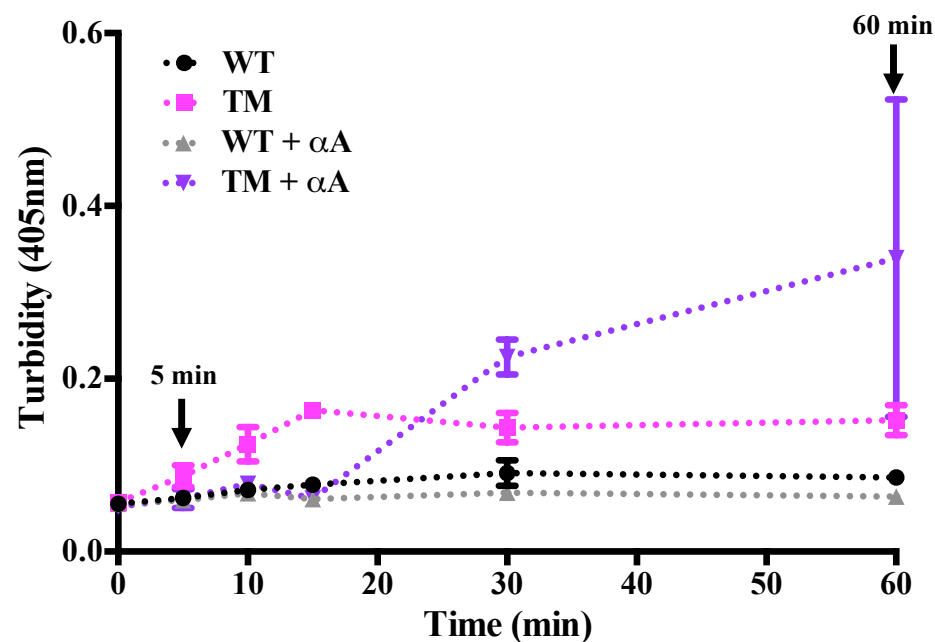
**Figure 7: Effect of deamidation on chemically-induced unfolding and aggregation of  $\gamma$ S.** WT and deamidated  $\gamma$ S at 5 mg/mL and in 2.2 M GuHCl were incubated in 96-microwell plates with shaking at 37 °C. **A.** Changes in turbidity at 405 nm. **B.** Changes in thioflavin T fluorescence at 440/485 nm were measured. Aggregation curves shown are average measurements ( $n \geq 2$ ) representative of three independent experiments. Inset shows maximum aggregation rate (Rate<sub>t50</sub>) derived from a Boltzmann sigmoid function ( $n \geq 4$ , error bars = SEM). Rate<sub>t50</sub> values that are significantly different from each other ( $p < 0.05$ ) are distinguished by different symbols.



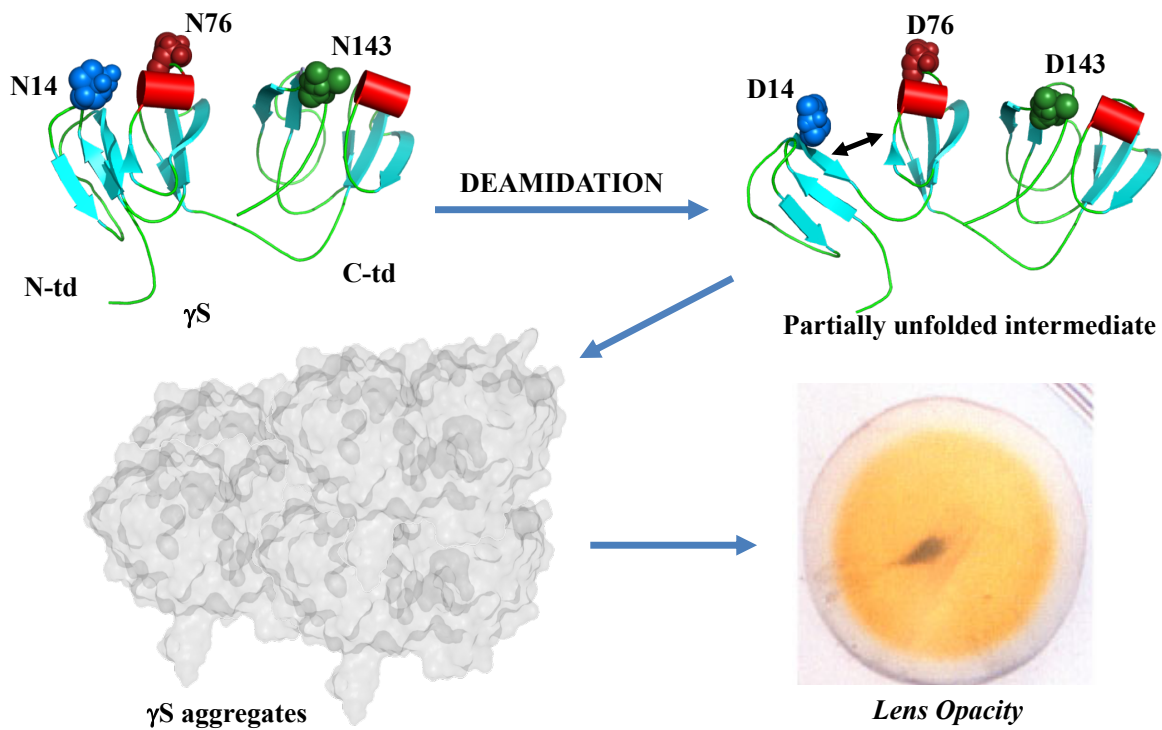
**B**

Rates of thermal-induced aggregation						
$\gamma$ S	Slope	p-value				
		WT	N14D	N76D	N143D	TM
WT	0.031	X	0.042	0.004	ns	0.006
N14D	0.065	0.042	X	ns	0.012	ns
N76D	0.060	0.004	ns	X	0.014	ns
N143D	0.039	ns	0.012	0.014	X	0.010
TM	0.102	0.006	ns	ns	0.010	X

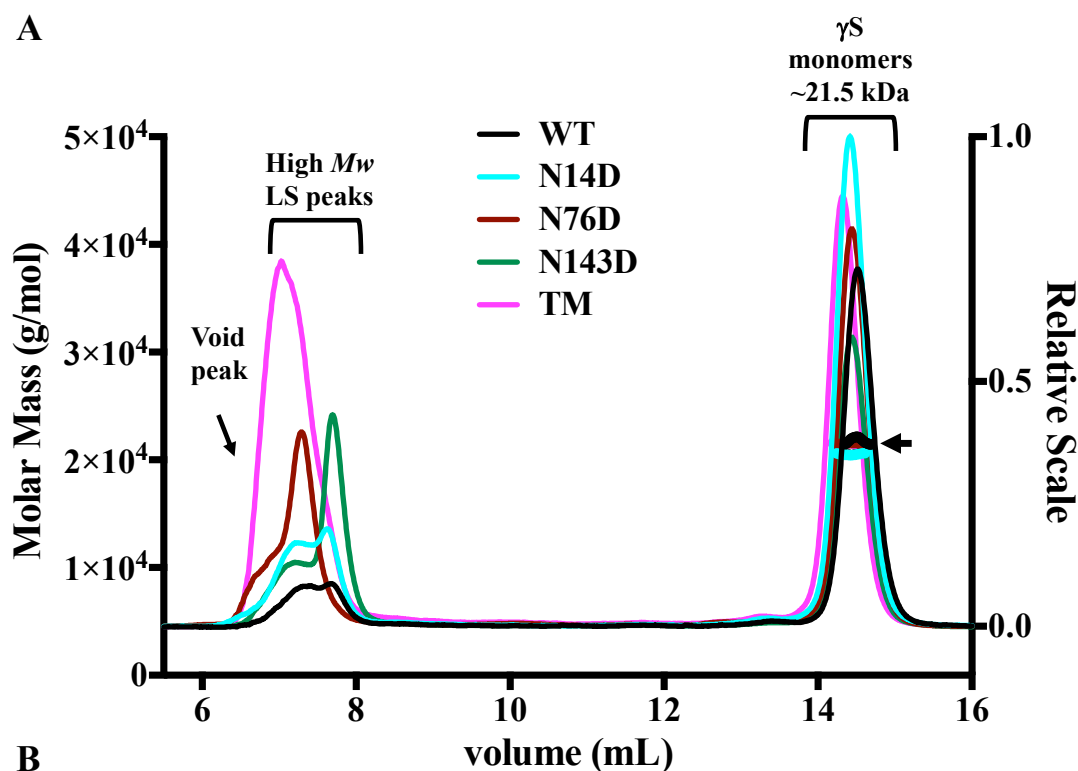
**Figure 8: Deamidation increases thermal-induced aggregation of  $\gamma$ S.** **A.** Changes in turbidity during thermal-induced denaturation of WT, N14D, N76D, N143D and TM. Proteins at 24  $\mu$ M concentrations were incubated at 70 °C and turbidity measured at 405 nm ( $n = 3$ , error bars = SEM). Note error bars are present on TM, but are not visible. Initial rates of aggregation were determined by linear regression. N14D, N76D and TM lines are significantly different than both WT and N143D. **B.** Table with p-values showing significant differences between aggregation rates for each  $\gamma$ S (ns = not significant).



**Figure 9: Thermal-induced aggregation of WT and TM  $\gamma$ S with and without  $\alpha$ A-crystallin.** WT and TM with and without  $\alpha$ A-crystallin were incubated at 65 °C and changes in turbidity at 405 nm were measured. (n = 3, error bars = SEM. Note: error bars are present on WT even though they are not visible at all time points and the large error bar on the TM at 60 min is due to sample precipitating.)

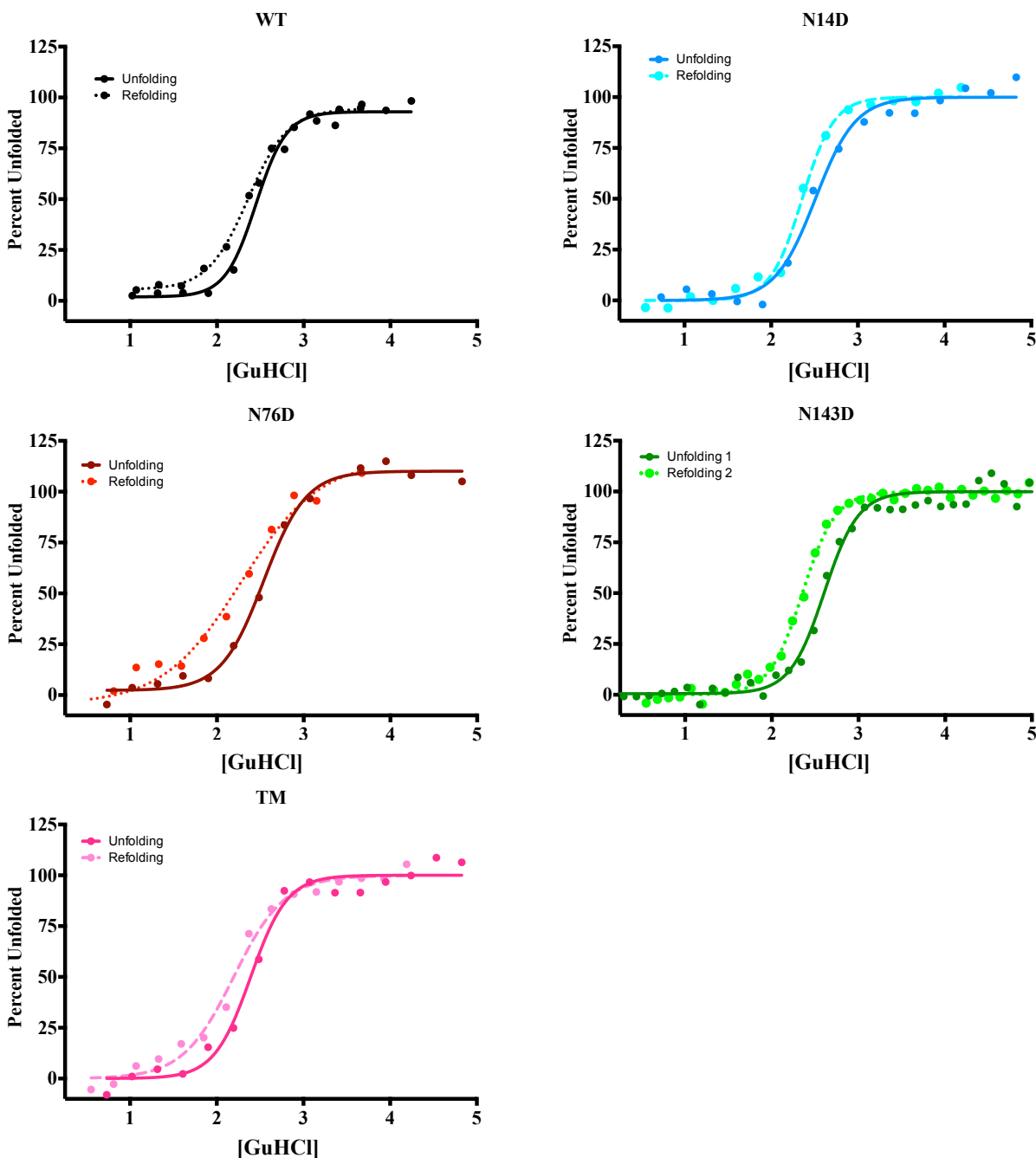


**Figure 10: Model summarizing the effects of deamidation on  $\gamma$ S.** Deamidations associated with cataracts increase propensity of  $\gamma$ S to partially unfold and form large light scattering aggregates that lead to lens opacity.



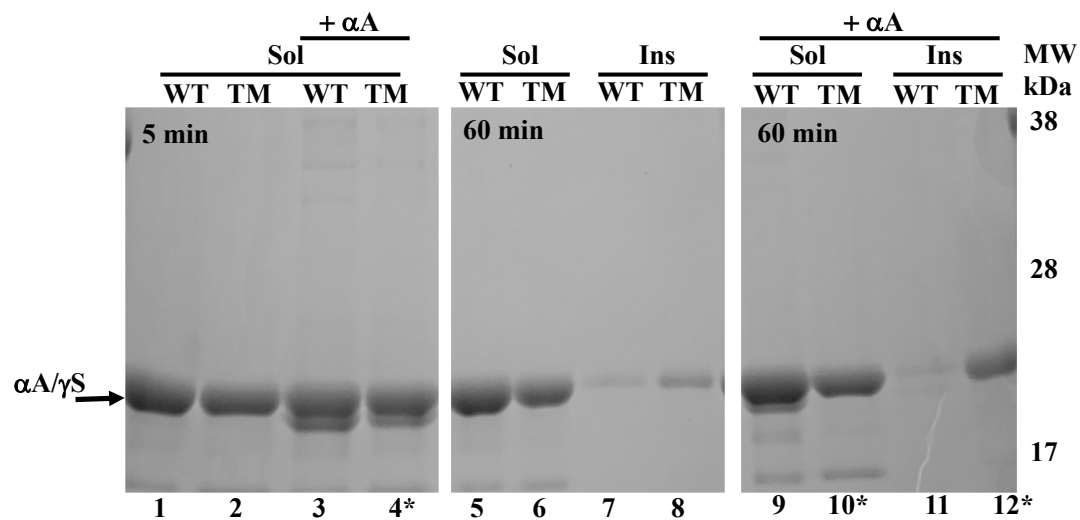
**Supplementary Figure 1: Increased relative light scattering of an aggregate  $\gamma$ S species due to deamidation.** **A.** A 50  $\mu$ L sample of 5 mg/mL WT, N14D, N76D, N143D and TM were separated by size exclusion chromatography in-line with a multiangle light scattering detector and a refractive index detector (Superose 6 column, GE Health, IL in-line with a Dawn Heleos-II and Optilab rEX, Wyatt Technology, CA). Samples were prepared in 58 mM sodium phosphate (pH 6.8), 100 mM KCl, 1 mM EDTA and 1 mM DTT. Between 140 and 240  $\mu$ g were recovered under the monomer peak eluting at 14.5 mL. The high molar mass peak eluted at 7.5 mL and was confirmed to contain  $\gamma$ S with mass spectrometry. **B.** Molar masses and polydispersity of  $\gamma$ S monomer from panel A. The molar mass of  $21.5 \times 10^3$  g/mol for all proteins is near the expected value of  $20.9 \times 10^3$  g/mol. The molar masses were not able to be determined for the aggregate, due to the low recovery of protein in this peak. Results were reproducible over multiple experiments.





**Supplementary Figure 2: Unfolding/refolding of WT and deamidated  $\gamma$ S showing hysteresis.**

Unfolding of deamidated mimics N14D, N76D, N143D, TM and WT  $\gamma$ S crystallin were incubated in water bath at 10  $\mu$ g/mL in 100 mM sodium phosphate, 1.1 mM EDTA, 5.5 mM DTT at 37°C for 12-16 hours with increasing amounts of guanidine hydrochloride (GuHCl). Refolding samples were unfolded as above, then diluted into decreasing concentrations of GuHCl and incubated at 37°C for 8-12 hours. Both sets of samples were then excited at 295 nm and emission spectra measured from 310-400 nm with slit widths set at 0.5 mm using fluorometer (Photon Technology Inc., NJ). The resulting 360 nm and 320 nm measurements were used to calculate percent protein unfolded at each GuHCl concentration and fitted to a sigmoidal two state unfolding curve using Prism 6 (GraphPad Software, CA).



**Supplementary Figure 3: Thermal-induced aggregation of WT and deamidated  $\gamma$ S with and without  $\alpha$ A.** SDS-PAGE of samples from 5 and 60 min of TM and WT incubated at 65 °C with and without  $\alpha$ A chaperone in 100 uL containing 100  $\mu$ g of protein with a ratio of 2:1,  $\gamma$ S: $\alpha$ A. Lane 1-4, soluble protein after 5 min of heating; Lanes 5-8, Soluble (Sol) and insoluble (Ins) proteins after 60 min of heating; Lanes 9-12, Soluble and insoluble protein after 60 min of heating with  $\alpha$ A. (\*) Bands from lanes 4, 10, and 12 were confirmed by mass spectrometry to contain  $\gamma$ S and  $\alpha$ A, which co-migrate.