Supporting Information

Banerjee et al. 10.1073/pnas.1014653107

SI Text

Appendix S1. A Mean-Field Model for Protein Mixtures. We use a mean-field thermodynamic analysis to further examine the connections between increased γ–α attraction, phase separation temperature, tie line direction, and light-scattering intensity. To do so we construct a mean-field model of the generalized van der Waals type (1), for the free energy of aqueous mixtures that are concentrated in both γ- and α-crystallin, and apply the mathematical conditions for phase separation and light scattering to consider the observed contrasts between E107A–α and HGD–α mixtures. The model is intended to illustrate the plausibility of the considerations above in a more mathematically specific manner, not to provide a quantitative model for the data.

The model is the sum of a part $k_BTf(\rho_1,\rho_2)$ that represents the mixing entropy of hard particles of the same size and shape as γ- and α-crystallins, and a part $u(\rho_1,\rho_2)$ that represents non hard-core interactions. $f(\rho_1,\rho_2)$ could be from an equation of state for mixtures of hard spheres of different sizes (2, 3), as in the unperturbed part of the first-order perturbation free energy used in (4), or by forms that also account for molecular shape (5). Here, we leave f unspecified in order to focus on the role of γ–α attractions.

In line with the regular solution approach to modeling liquid–liquid phase separation (6, 7), we take $u(\rho_1,\rho_2) = w_{\gamma\text{-}\gamma}\phi_1(1-\phi_1-\phi_2) + w_{\alpha\text{-}\alpha}\phi_2(1-\phi_1-\phi_2) + w_{\gamma\text{-}\alpha}\phi_1\phi_2$, in which $\phi_1 = \rho_1\bar{\nu}_1$ and $\phi_2 = \rho_2\bar{\nu}_2$ are the volume fractions of γ and α, respectively, and $\bar{\nu}_1$ and $\bar{\nu}_2$ denote their corresponding partial molecular volumes, which we take to be constant, together with that of the solvent. Note that $1-\phi_1-\phi_2$ is the volume fraction of solvent.

The quantities $w_{\gamma\text{-}\gamma}$, $w_{\alpha\text{-}\alpha}$ and $w_{\gamma\text{-}\alpha}$ represent average, effective attractions or repulsions between species and are proportional to the effective critical temperatures for liquid–liquid phase separation on the binary solvent-γ and solvent-α axes, and on a hypothetical γ–α axis without water. Further expressions that could be used to relate the w_{i-j} to the intermolecular potentials and the partial pair correlation functions are given in a perturbation theory model for γ–α mixtures (4), but there the composition dependences of the pair correlation functions are considered, and yield a model for $u(\rho_1,\rho_2)$ that better reflects the short-range of the interprotein attractions than the simpler form used here.

We take $\delta w_{\gamma\text{-}\alpha} = w_{\gamma\text{-}\alpha} - (w_{\gamma\text{-}\gamma} + w_{\alpha\text{-}\alpha})$ to represent effective solvent-mediated γ–α interactions; note that increased γ-α attraction decreases $w_{\gamma\text{-}\alpha}$ and hence $\delta w_{\gamma\text{-}\alpha}$. We take $w_{\alpha\text{-}\alpha} = 0$ because α-crystallin solutions are modeled well by the hard-sphere Carnahan–Starling equation of state (8, 9) and at physiological pH α-crystallin solutions do not phase separate; see for example refs. 9 and 10. Thus we model

$$\begin{aligned} g &= G/Vk_BT \\ &= f(\rho_1,\rho_2) + (w_{\gamma\text{-}\gamma}/k_BT)\phi_1(1-\phi_1) + (\delta w_{\gamma\text{-}\alpha}/k_BT)\phi_1\phi_2. \end{aligned} \quad \textbf{[S1]}$$

In Eq. **S1** g is the Gibbs free energy G per unit volume V, per unit thermal energy k_BT.

To apply the model, we note that the present data show the critical temperatures and therefore the $w_{\gamma\text{-}\gamma}$ values for E107A-solvent and HGD-solvent solutions to be nearly identical. Second, the data show E107A and HGD to be similar in folding, so f will also be identical for E107A–α and HGD–α mixtures. Thus, differences between the mixtures will be modeled solely by differing $\delta w_{\gamma\text{-}\alpha}$, or equivalently by changes in $w_{\gamma\text{-}\alpha}$, given our assumptions.

In the following we take g_{ijk} to denote a mixed partial derivative of g that is taken i times with respect to ρ_1, j times with respect to ρ_2, and k times with respect to T. We treat the buffer as a single component and assume that all partial molecular volumes are constant. It is convenient to note that under these assumptions differentiation with respect to ρ_1, while holding ρ_2 fixed (with solvent volume fraction remaining $1-\phi_1-\phi_2$) gives a result that, when multiplied by $(1/\bar{\nu}_1)$, is the same as that from differentiation with respect to ϕ_1, while holding ϕ_2 fixed.

The spinodal boundary between unstable and metastable regions satisfies

$$g_{200}g_{020} - (g_{110})^2 = 0. \quad \textbf{[S2]}$$

Substituting Eq. **S1** into Eq. **S2**, we obtain

$$(f_{200} - (2\bar{\nu}_1^2 w_{\gamma\text{-}\gamma}/k_BT))(f_{020}) - (f_{110} + (\bar{\nu}_1\bar{\nu}_2\delta w_{\gamma\text{-}\alpha}/k_BT))^2 = 0. \quad \textbf{[S3]}$$

The left-hand side of Eq. **S3** is quadratic in $\delta w_{\gamma\text{-}\alpha}$ and thus clearly nonmonotonic in that quantity. Because $w_{\gamma\text{-}\gamma}$ is the same for E107A and HGD, and at a given composition the derivatives of f are also the same for the E107A–α and HGD–α systems, Eq. **S3** is qualitatively consistent with the findings of Dorsaz et al. (4) that the stability of the mixtures is expected to be nonmonotonic in the strength of effective, solvent-mediated heterologous, γ–α attractions, $\delta w_{\gamma\text{-}\alpha}$. This follows because $\delta w_{\gamma\text{-}\alpha}$ values that are too negative or too positive decrease the left-hand side and therefore enhance instability.

Such nonmonotonic dependence of stability on free energy parameters is well-known from other ternary mixture contexts, for example in the case of the regular solution model (6). To understand this implication further, note that in the term in Eq. **S2** from which the present nonmonotonic dependence originates, $-(g_{110})^2$, the quantity g_{110} represents the coefficient of a saddle-like term in the free energy, that has concave down directions that promote instability for either sign of nonzero values of g_{110}; broadly speaking, here these directions correspond to compositional and density-type phase separation.

Eq. **S2** implies that there is a composition direction along which the second directional derivative of the free energy vanishes, that of the eigenvalue *0* eigenvector of the Hessian $H_\rho[g] = \begin{pmatrix} g_{200} & g_{110} \\ g_{110} & g_{020} \end{pmatrix}$. That vector is perpendicular to both row vectors of H and therefore has slope $(d\rho_2/d\rho_1)_{sp,0} = -g_{110}/g_{020} = -g_{200}/g_{110}$.

The direction $(d\rho_2/d\rho_1)_{sp,0}$ has special significance at a critical point because its value there $(d\rho_2/d\rho_1)_{cr,0}$ is that to which nearby, short, tie line slopes are asymptotic. Thus, through this connection, the observed directions of the tie lines near such a critical point can be taken to estimate the value of g_{110}/g_{020} there. Note that in Fig. S4*II*, by including the measured cloud point temperatures for both of the $X = 0.5$ mixtures, indicates that both observed tie lines must indeed lie rather close though somewhat below corresponding critical loci, so that considerations of this paragraph are expected to apply.

More specifically, for the present model we find that

$$(d\rho_2/d\rho_1)_{sp,0} = -(f_{110} + (\bar{\nu}_1\bar{\nu}_2\delta w_{\gamma\text{-}\alpha}/k_BT))/f_{020}. \quad \textbf{[S4]}$$

Eq. **S4** suggests that increasing γ–α attractive interactions and therefore decreasing $\delta w_{\gamma\text{-}\alpha}$ will tend to increase asymptotic tie line slopes near a critical point in the ternary concentration triangle, consistent with the data on E107A–α and HGD–α mixtures shown in Fig. 4. However, it should be noted that the connection between $\delta w_{\gamma\text{-}\alpha}$ and tie line slopes near a critical point is not direct, because changing $\delta w_{\gamma\text{-}\alpha}$ can change critical point locations and thereby change relevant values of f_{110} and f_{020} in Eq. **S4**. From a physical point of view, the rotation of tie lines toward density-type phase separation can be understood by considering this phenomenon as a much less extreme version of a hypothetical case of stable complex formation due to attractive interactions between the two dissimilar proteins, in which case the tie lines would rotate still further so as to nearly become rays through the pure solvent vertex of the composition triangle.

The Rayleigh ratio in excess of solvent, ΔR, can also be expressed in terms of g, for small scattering vector magnitudes (10, 11). For constant partial specific volumes, one can show (11)

$$\Delta R = (\pi^2/\lambda^4)\nabla_\rho \varepsilon \cdot H_\rho[g]^{-1} \cdot \nabla_\rho \varepsilon \qquad \textbf{[S5]}$$

where $\nabla_\rho \varepsilon$ is the gradient $(\epsilon_{100}, \epsilon_{010})$ of the dielectric constant ϵ, $H_\rho[g]^{-1}$ is the inverse of the Hessian of g with respect to ρ_1 and ρ_2, and λ is the wavelength in vacuum of the incident light, assumed polarized normal to the scattering plane, as for the present data. Briefly, Eq. **S5** implies that small values of the directional second composition derivatives of g increase light scattering intensity, the more so when they are aligned with the dielectric constant gradient vector; a precise formulation of this statement may be found in (11). Using Eq. **S1** in Eq. **S5** we obtain

$$\begin{aligned}\Delta R &= (\pi^2/\lambda^4)[\{\epsilon_{100}^2, \epsilon_{010}^2, -2\epsilon_{100}\epsilon_{010}\} \\ &\quad \cdot \{g_{200}, g_{020}, g_{110}\}/(g_{200}g_{020} - (g_{110})^2)] \\ &= (\pi^2/\lambda^4)[\{\epsilon_{100}^2, \epsilon_{010}^2, -2\epsilon_{100}\epsilon_{010}\} \\ &\quad \cdot \{f_{200} - (2\bar{\nu}_1{}^2 w_{\gamma\text{-}\gamma}/k_B T), f_{020}, f_{110} \\ &\qquad + (\bar{\nu}_1\bar{\nu}_2\delta w_{\gamma\text{-}\alpha}/k_B T)\}/(g_{200}g_{020} - (g_{110})^2)]. \qquad \textbf{[S6]}\end{aligned}$$

We now consider the implications of Eq. **S6** for the contrast between E107A–α and HGD–α mixtures. First, at a given temperature and composition in the single phase region, the closer HGD–α spinodal suggests that the right-hand side denominator will in general be smaller for HGD-α mixtures. Therefore the observed increase in light scattering of E107A–α mixtures suggests that the corresponding numerator of ΔR would now be larger for E107A–α mixtures than for HGD–α mixtures. To see what can be inferred for $\delta w_{\gamma\text{-}\alpha}$, note first the dot product in the numerator and that ϵ_{100} and ϵ_{010} are positive. Becuase $w_{\gamma\text{-}\gamma}$ is the same for E107A and HGD, increased light scattering intensity observed for E107A–α mixtures implies that $\delta w_{\gamma\text{-}\alpha}$ is smaller for E107A–α than for HGD–α mixtures, again consistent with increased γ–α attractions in E107A–α mixtures.

In summary, the model is consistent with (a) nonmonotonic dependence of phase separation temperature on $\delta w_{\gamma\text{-}\alpha}$ (Eq. **S3**), (b) rotation of tie lines toward density–density phase separation implying increased γ–α attractions (Eq. **S4**), and (c) linkage between increased γ–α attractions and the observed, increased light scattering intensity (Eq. **S6**).

However, in light of (*i*) the well-established role of short-range attractions in determining the phase boundaries of protein solutions (12), (*ii*) the success of molecular dynamics and perturbation theory that include short-range attraction in modeling neutron scattering data for γ- and α-crystallin mixtures (4, 13), together with (*iii*) the importance of aeolotopic, or anisotropic interactions in affecting the single-protein component phase boundaries (14), it is important to note that the present model assumes functional forms for free energy that need modification in order to account for short-range, aeolotopic attractions. It can nevertheless be a guide for expectations as to the phenomena that can result from increased γ–α attractions.

Appendix S2. A Computational Protein Model for Attractive Interactions Between HGD/E107A and Human αB-crystallin (HAB). Electrostatic surface potential was calculated using the Delphi software package, v.4 (15) for HGD, E107A and HAB as shown in Fig. S5. The E107A structure was derived from the HGD structure (16) (PDB ID: 1hk0), by making the Ala substitution at position 107, followed by energy minimization using the DeepView PDB viewer (17). The HAB structure was modeled using the recently reported structure of truncated HAB (18). Even though bovine α-crystallin is an oligomer of the αA- and αB-crystallin chains present in a 3:1 ratio, αB-crystallin is known to decorate the surface of the oligomer while αA-crystallin predominantly resides within the oligomer core (19, 20). Therefore, we have chosen to use HAB as the representative of α-crystallin in these studies. Thus the choice of the HAB structure instead of the bovine protein is based on the availability of the recent structure of the human protein. This choice is reasonable due to the overall strong sequence identity (which exceeds 94% between the two proteins), with a total identity in the residues of direct interest for this study. The interaction of HGD and E107A with HAB was investigated using the AUTODOCK program (21). Pymol graphics software (22) was used to map the electrostatic potential on the protein surface.

Fig. S5 shows the electrostatic potential mapped on the surface of HGD (Fig. S5.*A*) and E107A (Fig. S5.*B*). The surface of E107A shows a distinct positive-potential patch (blue patch around Ala 107) which also includes arginines 89, 115, and 169 and which can interact effectively with the negatively charged α-crystallin. A similar patch in HGD (Fig. S5*A*) includes these arginines as well, but the negatively charged Glu107 in HGD disrupts the positive-potential.

Fig. S5 also shows the electrostatic potential on the surface of HAB (PDB ID: 3L1G) with a negative-potential patch in red, that is likely to show attractive interactions with the opposite potential on the surface of HGD or E107A. The prominent large negative patch on the surface of HAB (Fig. S5*C*) includes residues Glu87, 88, 105, 106, Asp109, and Glu110. We assume this large negative-potential patch as the likely site of interaction between HAB and HGD/E107A. Although there are other negative-potential patches on the surface of HAB, they are much smaller in size, and have not been modeled in this work.

The surface electrostatic potential maps on these proteins (Fig. S5) guided us in deciding the sites of interaction for our modeling studies. These studies were carried out to obtain a working structural model to estimate changes in the attractive electrostatic interaction energies for the HGD–α and E107A–α interactions. We used the AUTODOCK program (21), which is widely used to "dock" a protein with another protein or ligand. The program also estimates the free energies of interaction for the best fit in the docking experiments. We set the γ-crystallin (either HGD or E107A) as the "protein" and HAB as the "ligand." This selection enabled us to set the grid box dimensions (which specify the binding site on the protein) in such a way that it included the positive patch on E107A, and an identical location on HGD but without the positive patch. All the residues of HGD or E107A were considered to be rigid. For the ligand, namely HAB, we used a six-residue HAB fragment, which contained residues Glu105 to Glu110 extracted from the HAB crystal structure (18), and which represented the negatively charged patch on HAB. The docked conformations for HGD–α and E107A–α are shown in Fig. S6 *A* and *B*, respectively. Alternate views of the docking results are presented in Figs. S6 *C* and *D* that clearly highlight the greater overlap of Arg 115 in E107A with Glu105

in HAB (Fig. S6*D*), which is not the case in HGD (Fig. S6*C*). This best fit configuration yields an electrostatic interaction energy of −2.04 kcal/mol for the HGD–HAB interaction and −2.43 kcal/mol for the E107A–HAB interaction, representing a change in interaction energy that is consistent with stronger binding of E107A with α-crystallin.

1. Barker JA, Henderson D (1976) What is a liquid—understanding states of matter. *Rev Mod Phys* 48:587–671.
2. Boublik T (1970) Hard-sphere equation of state. *J Chem Phys* 53:471–472.
3. Mansoori GA, Carnahan NF, Starling KE, Leland TW (1971) Equilibrium thermodynamic properties of mixture of hard spheres. *J Chem Phys* 54:1523–1525.
4. Dorsaz N, Thurston GM, Stradner A, Schurtenberger P, Foffi G (2009) Colloidal characterization and thermodynamic stability of binary eye lens protein mixtures. *J Phys Chem B* 113:1693–1709.
5. Gibbons RM (1970) Scaled particle theory for mixtures of hard convex particles. *Mol Phys* 18:809–816.
6. Meijering JL (1950) Segregation in regular ternary solutions. 1. *Philips Res Rep* 5:333–356.
7. Furman D, Dattagupta S, Griffiths RB (1977) Global phase-diagram for a 3-component model. *Phys Rev B* 15:441–464.
8. Carnahan NF, Starling KE (1969) Equation of state for nonattracting rigid spheres. *J Chem Phys* 51:635–636.
9. Veretout F, Delaye M, Tardieu A (1989) Molecular basis of eye lens transparency. Osmotic pressure and X-ray analysis of alpha-crystallin solutions. *J Mol Biol* 205:713–728.
10. Thurston GM (2006) Liquid-liquid phase separation and static light scattering of concentrated ternary mixtures of bovine alpha and gammaB crystallins. *J Chem Phys* 124:134909.
11. Ross DS, Thurston GM, Lutzer CV (2008) On a partial differential equation method for determining the free energies and coexisting phase compositions of ternary mixtures from light scattering data. *J Chem Phys* 129:064106.
12. Lomakin A, Asherie N, Benedek GB (1996) Monte Carlo study of phase separation in aqueous protein solutions. *J Chem Phys* 104:1646–1656.
13. Stradner A, Foffi G, Dorsaz N, Thurston G, Schurtenberger P (2007) New insight into cataract formation: Enhanced stability through mutual attraction. *Phys Rev Lett* 99:198103.
14. Lomakin A, Asherie N, Benedek GB (1999) Aeolotopic interactions of globular proteins. *Proc Natl Acad Sci USA* 96:9465–9468.
15. Rocchia W, Alexov E, Honig B (2001) Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *J Phys Chem B* 105:6507–6514.
16. Basak A, et al. (2003) High-resolution X-ray crystal structures of human gammaD crystallin (1.25 A) and the R58H mutant (1.15 A) associated with aculeiform cataract. *J Mol Biol* 328:1137–1147.
17. Kaplan W, Littlejohn TG (2001) Swiss-PDB viewer (Deep View). *Brief Bioinform* 2: 195–197.
18. Laganowsky A, et al. (2010) Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci* 19:1031–1043.
19. Asselbergs FA, Koopmans M, Van Venrooij WJ, Bloemendal H (1978) Posttranslational assembly of lens alpha-crystallin in the reticulocyte lysate and in Xenopus laevis oocytes. *Eur J Biochem* 91:65–72.
20. Aquilina JA, et al. (2005) Subunit exchange of polydisperse proteins: Mass spectrometry reveals consequences of alphaA-crystallin truncation. *J Biol Chem* 280: 14485–14491.
21. Morris GM, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19:1639–1662.
22. DeLano WL (2008) The Pymol Molecular Graphics System.

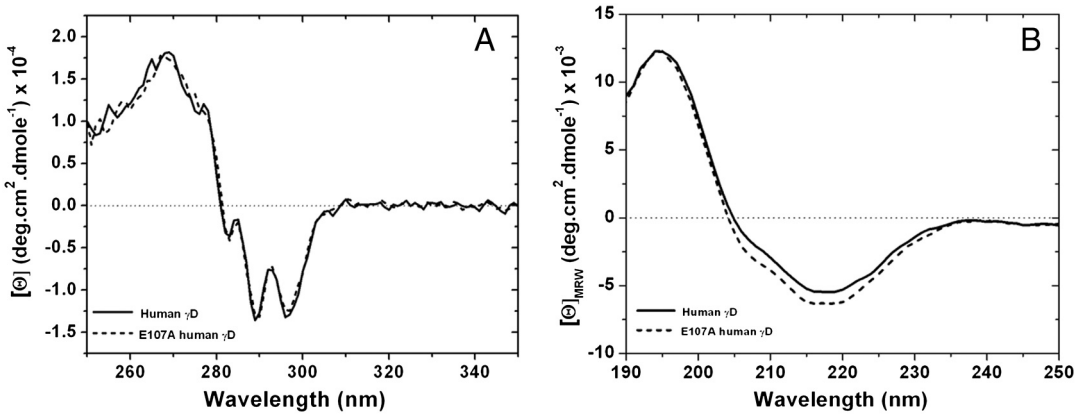

Fig. S1. (*A*) Near-UV CD spectra of HGD and E107A in 0.1 M phosphate buffer, pH 7.0. Protein concentration, 0.5 mg/mL. (*B*) Far-UV CD spectra of HGD and E107A in 5 mM phosphate buffer, pH 7.0. Protein concentration, 0.1 mg/mL.

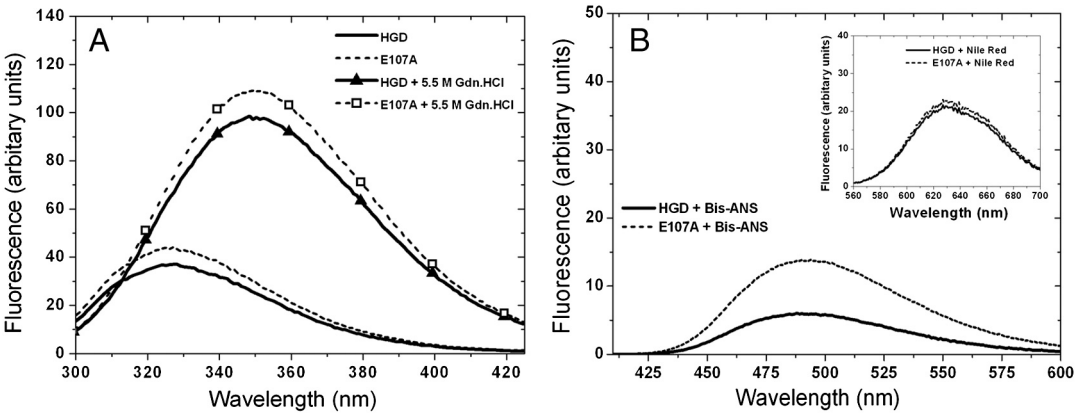

Fig. S2. (*A*) Tryptophan fluorescence emission spectra of folded and unfolded HGD and E107A (excitation at 290 nm) in 0.1 M phosphate buffer, pH 7.0, protein concentration, 0.02 mg/mL. Folded HGD (—), E107A (·····); unfolded with guanidinium hydrochloride (Gdn.HCl), HGD (-▲-), E107A (-□-). (*B*) Fluorescence spectra of HGD (—) and E107A (·····) with Bis-ANS (excitation at 390 nm) in 0.1 M phosphate buffer, pH 7.0, protein concentration, 0.1 mg/mL, and Bis-ANS concentration, 100 μM. (*Inset*) Fluorescence spectra of HGD and E107A with Nile Red in 0.1 M sodium phosphate buffer, pH 7; excitation, 540 nm, protein concentration, 0.1 mg/mL, and Nile Red concentration, 100 μM.

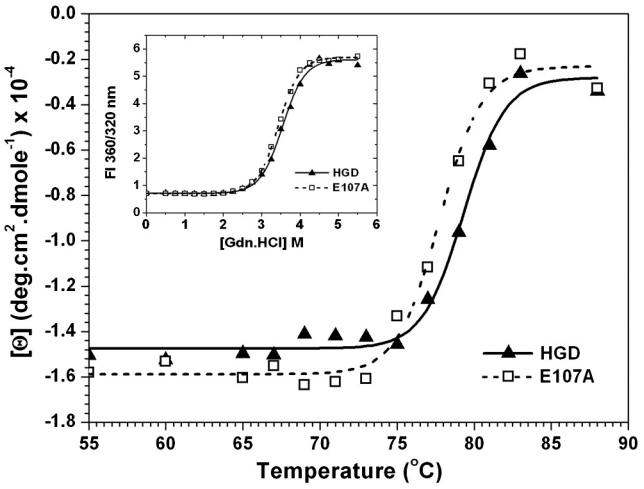

Fig. S3. Thermal stability profiles for HGD and E107A in 0.1 M sodium phosphate buffer, pH 7.0, measured as change in ellipticity at 290 nm. (*Inset*) Gdn.HCl-induced unfolding of HGD and E107A in 0.1 M sodium phosphate buffer, pH 7.0. Unfolding monitored by the ratio of the emission intensities at 320 nm and 360 nm, with excitation at 290 nm.

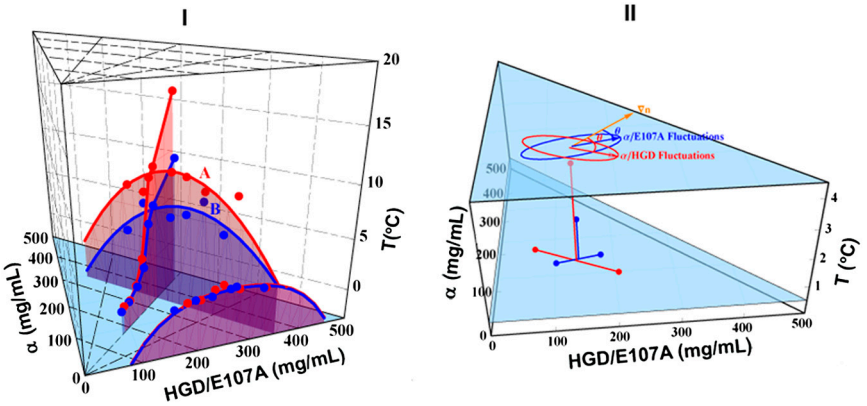

Fig. S4. (*I*) Mixtures of E107A and α-crystallin (dark blue, B) have lower cloud point temperatures than mixtures of HGD and α-crystallin (red, A), while solutions of pure HGD and E107A have nearly identical coexistence curves (purple). Assembled coexistence curves are plotted above the ternary α/γ/buffer concentration triangle. (*II*) A 3D view of the direction of the measured tie lines (on lower blue triangle) and the differing principal fluctuation directions they suggest (on upper blue triangle) for the nearby, single phase region. E107A/α-crystallin fluctuations are projected to be much more closely aligned with the gradient vector of refractive index than those for HGD/α-crystallin mixtures at $X_a = 0.50$ (see text). The tops of the vertical red and blue lines denote the T_{ph} of HGD–α and E107A–α mixtures, respectively.

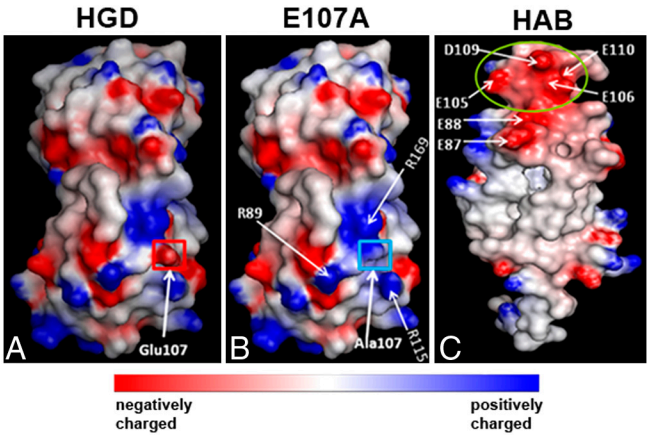

Fig. S5. Electrostatic potential mapped on the structures of (*A*) HGD, (*B*) E107A, and (*C*) Human αB-crystallin (HAB). The corresponding color scale is shown at the bottom. The mutation site in HGD is shown in the boxed region (Glu107 to Ala). A large part of the negative-potential patch on the surface of HAB which was used for docking studies has been circled in green. See text for details.

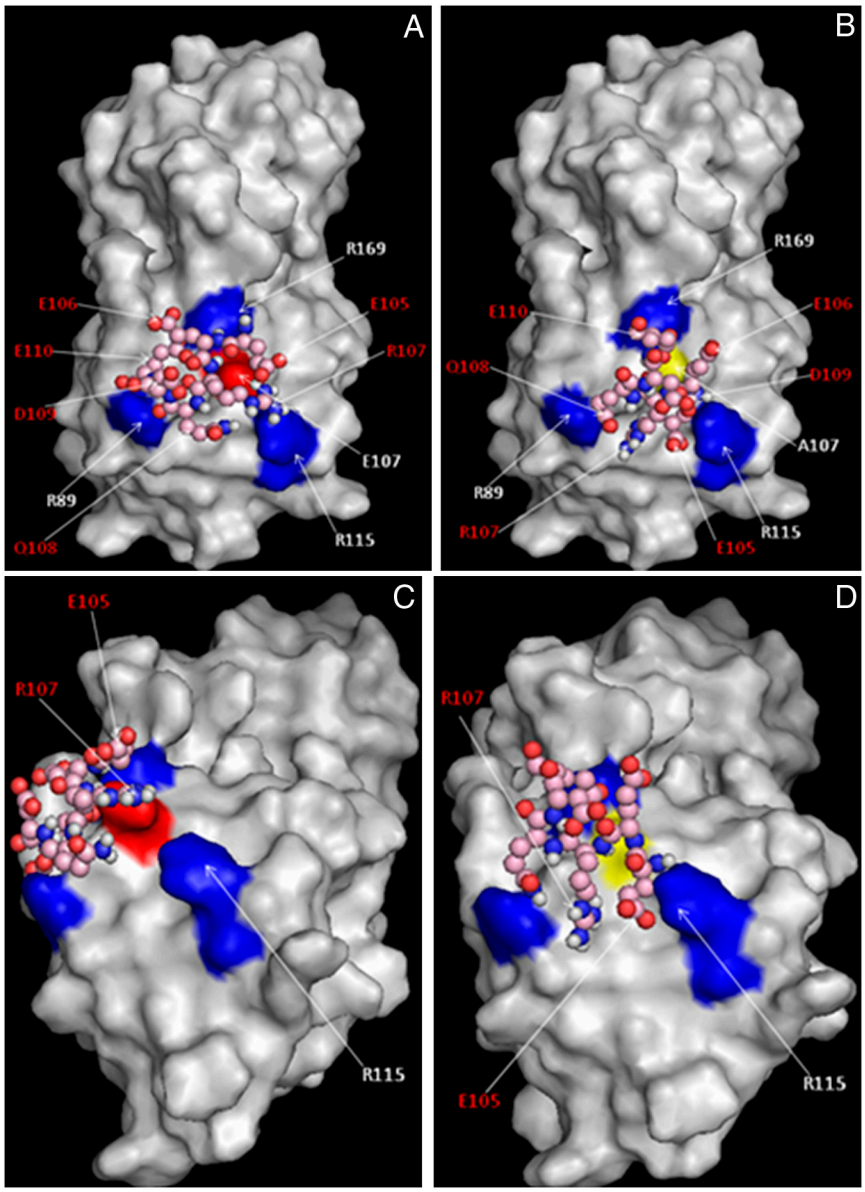

Fig. S6. Docked conformations showing a model peptide fragment of HAB interacting with (*A*) HGD and (*B*) E107A. HAB residues are labeled in red and HGD/E107A residues are labeled in white. The mutation site in E107A is shown in yellow. Different views of the same conformations are shown in *C* and *D* to stress that Arg115 in HGD is not interacting with the HAB peptide (*C*), while in E107A the same residue is interacting with Glu105 of the HAB peptide (*D*).