Altered phase diagram due to a single point mutation in human γD-crystallin

Jennifer J. McManus*, Aleksey Lomakin*, Olutayo Ogun*, Ajay Pande*†, Markus Basan‡, Jayanti Pande†‡, and George B. Benedek*‡§¶

*Materials Processing Center, ‡Department of Physics, and §Center for Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

Contributed by George B. Benedek, August 7, 2007 (sent for review June 27, 2007)

The P23T mutant of human γD-crystallin (HGD) is associated with cataract. We have previously investigated the solution properties of this mutant, as well as those of the closely related P23V and P23S mutants, and shown that although mutations at site 23 of HGD do not produce a significant structural change in the protein, they nevertheless profoundly alter the solubility of the protein. Remarkably, the solubility of the mutants decreases with increasing temperature, in sharp contrast to the behavior of the native protein. This inverted solubility corresponds to a strong increase in the binding energy with temperature. Here we have investigated the liquid–liquid coexistence curve and the diffusivity of the P23V mutant and find that these solution properties are unaffected by the mutation. This means that the chemical potentials in the solution phase are essentially unaltered. The apparent discrepancy between the interaction energies in the solution phase, as compared with the solid phase, is explicable in terms of highly anisotropic interprotein interactions, which are averaged out in the solution phase but are fully engaged in the solid phase.

cataract | lens | protein phase diagram | quasielastic light scattering

Human γD-crystallin (HGD) is an important member of the γ-crystallin family of proteins found in the human lens. Mutations in HGD in particular have been associated with a number of childhood cataracts (1–4). Recently, the P23T mutant of HGD has been associated with coralliform, cerulean, and fasciculiform cataract phenotypes (refs. 1 and 2 and references therein). Mutation at site 23 of HGD is only one of a number of single point mutations, including the R14C, R58H, and R36S mutations, that occur on the CRGD gene and that have been linked with early-onset cataract disease (5–9). Physicochemical characterization of the mutant proteins shows why these changes result in the formation of either covalently linked aggregates in the case of R14C (5, 6) or crystals in the case of the R58H and R36S mutants (7–9). The P23T mutation results in decreased solubility of the protein, leading to protein aggregation and light scattering, and hence to lens opacity. However, the aggregates are not covalently linked, since the aggregation process is completely reversible with temperature.

Formation of protein aggregates is a common motif in many "condensation diseases," which include cataract (10), sickle-cell anemia (11, 12), and Alzheimer's disease (13, 14), as well as other amyloid diseases such as diabetes and Parkinson's disease (15, 16). These examples highlight the importance of understanding the processes that lead to the formation of condensed protein phases under physiological conditions. In particular, protein condensation diseases resulting from a single amino acid substitution provide favorable conditions for biophysical analysis because such substitutions produce a change in the interaction potential between the proteins but may not necessarily result in significant structural changes to the protein itself. The first example of a protein condensation disease being associated with a single amino acid substitution was sickle cell anemia (11). Interestingly, hemoglobin S (HbS) displays retrograde solubility in its phase diagram and also a lower consolute temperature (17).

A more detailed investigation of P23T and other mutants of HGD at site 23 (i.e., P23S and P23V) has shown that unusual phase behavior is displayed by all three of these mutants, implying that the region around site 23 of HGD is important in maintaining the solubility of the protein (1). Although substituting a single amino acid residue at site 23 does not alter the regular form of the coexistence curve, it results in decreased protein solubility and remarkably inverts the form of the solubility line. The mutation responsible for the inversion of the solubility line *in vivo* is P23T. Unfortunately, P23T has very low solubility [<5 mg/ml; $\phi = 0.0035$. The volume fraction $\phi = c \cdot \upsilon_{sp}$, where c is the concentration in milligrams per milliliter and υ_{sp} is the partial specific volume 7.1×10^{-4} ml/mg (18) over the temperature range investigated]. With such low solubility, many features of the phase diagram are not accessible and cannot be measured, including liquid–liquid phase separation (LLPS). We chose to focus on the more soluble P23V substitution because this mutant demonstrates the same inverse solubility as P23T but over a larger concentration range, allowing us to investigate LLPS, aggregation, and crystallization. Although we will focus our experimental efforts on the P23V mutant, it will be possible to comment on the behavior of all three mutants.

We first summarize the state of our knowledge regarding the P23V mutant of HGD. The solubility line of mutants of HGD at site 23 shows a retrograde temperature dependence. The replacement of Thr-23 with a Ser or Val residue shifts the location of the inverted solubility line to higher concentrations (1). On the other hand, LLPS for P23V appears to be the same as for HGD, and the CD, Raman, and IR spectra show that no major secondary or tertiary structural changes take place upon mutagenesis (1).

Given the unusual nature of the phase diagram, we undertook further studies of these phase separation phenomena. We have extended the data for the LLPS, confirming that it is in fact the same as for the native protein. Furthermore, we have used quasielastic light scattering (QLS) spectroscopy to probe the interactions between the proteins in the soluble region of the phase diagram. We have also produced crystals of P23V and will comment on this in relation to the phase diagram.

Results and Discussion

Phase Diagram. The P23V phase diagram is shown in Fig. 1. Here we redraw the solubility data from ref. 1 and include our

Author contributions: J.J.M., J.P., and G.B.B. designed research; J.J.M., A.L., O.O., A.P., and M.B. performed research; J.J.M., A.L., and G.B.B. analyzed data; and J.J.M., A.L., and G.B.B. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: HbS, hemoglobin S; HGD, human γD-crystallin; LLPS, liquid–liquid phase separation; QLS, quasielastic light scattering.

†Present address: Department of Chemistry, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY 12222.

¶To whom correspondence should be addressed at: Room 13-2005, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. E-mail: benedek@mit.edu.

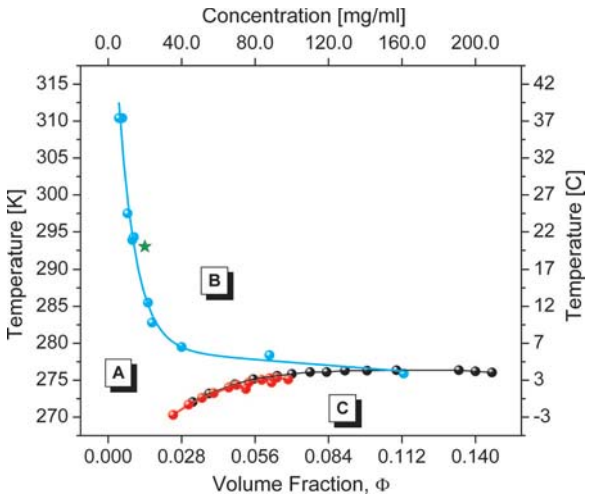

Fig. 1. Phase diagram for P23V, showing the coexistence curve for HGD (black spheres), the partial coexistence curve for P23V (red spheres), and the solubility line for P23V (blue spheres). The green star is a single point of the liquidus line, representing crystal–monomer equilibrium. The solubility line is replotted from numerical values given in ref. 1. In region A, the protein exists as a homogenous monomer. In region B, the protein is aggregated and is coexisting with the protein monomer (see lower part of Fig. 2). In region C, LLPS has occurred. The protein exists in two coexisting liquid phases, one concentrated and one dilute. Both liquid phases contain monomeric protein.

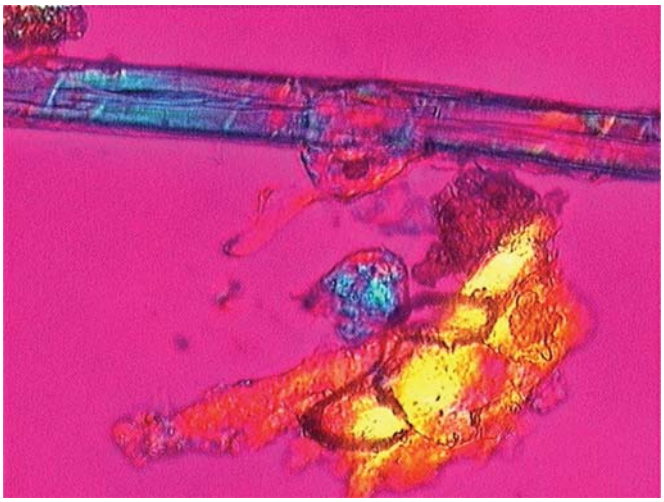

Fig. 2. Polarization microscopy image of P23V crystals. The uneven edges in the lower part of the image are due to melting. Crystals were grown at 32°C and then imaged at room temperature.

measured coexistence curve for the protein. Only a portion of the LLPS phase boundary is shown. On the phase diagram, there is a narrow gap between the coexistence curve and the apparent solubility line in region A where the protein can be brought to higher ϕ without aggregation. The maximum protein concentration we could reach at a temperature of 277 K (4°C) was ≈100 mg/ml (ϕ = 0.071) before aggregation began to occur. Consequently, this was the highest concentration at which measurements of LLPS could be taken. The LLPS for the native protein is also shown for comparison. Clearly, the coexistence curves for both native and mutant proteins are the same.

In region B of the phase diagram, the protein condenses to form a solid phase that coexists with the protein monomer. The nature of the condensed, solid phase depends on the preparation conditions of the protein solution. In the vast majority of region B, the solid phase exists as protein aggregates. In a very narrow part of region B, we have been able to suppress aggregation long enough to allow crystallization to occur. The line that divides regions A and B of the phase diagram, which we call the "solubility line," describes the coexistence of protein aggregates with protein monomer. This protein aggregation is completely reversible with temperature and, on the time scale of our experiments, appears to be at equilibrium. Therefore, for this work we assume that the solubility line is an equilibrium phase boundary.

We produced crystals of P23V within a narrow section of region B on the phase diagram, when solutions were passed through 0.02-μm filters (Fig. 1). Samples at concentrations between 20 and 40 mg/ml (ϕ = 0.014 to 0.028) were incubated at 32°C (305 K) overnight. Visual examination of the glass tubes revealed that small crystals had formed on the walls; their crystalline order was confirmed by observation of optical birefringence (Fig. 2). These crystals were of very low quality and only grew in small numbers. Many were irregularly shaped. Attempts to improve the number and quality of the crystals failed because aggregation often occurred when we tried to manipulate the samples. This further highlighted the competition between crystallization and aggregation. Protein crystals produced from a solution at a concentration of 20 mg/ml (ϕ = 0.014) began to melt at room temperature (20°C, 293 K), as can be seen at the edges of the protein crystal in Fig. 2. To establish a point on the liquidus line, crystals are usually placed in a small volume of protein-free solvent. The protein crystals will dissolve until equilibrium has been reached. Because of the small number of crystals that grew, we determined the solubility of the protein crystal by starting from a supersaturated solution (19, 20). In solutions in which crystals grew, we lowered the temperature until melting began to occur at the edges of the crystals. We took this as the solubility. When this method is used, the surface may be poisoned by impurities or imperfections arising from improperly oriented proteins, thereby stopping further crystal growth (21, 22); however, previous work (20) has shown that in practice there is no significant difference in the results obtained using either method.

The point at which crystal melting occurs upon lowering of the temperature is shown on Fig. 1 (green star). This point lies very close to the solubility line (defined by blue spheres in Fig. 1). Because crystals form very close to this solubility line, we can assume that aggregation occurs at low supersaturation and, therefore, that the solubility line lies very close to the liquidus line that describes equilibrium between protein crystals and protein monomer.

Because aggregation occurs at low supersaturations, we found only a narrow window in the phase diagram in which aggregation could be suppressed in favor of crystallization. Essentially, we suppressed aggregation long enough to allow crystallization to occur. Given that the crystals grew only on the walls and the surface of the protein solutions, this crystal growth was by heterogeneous nucleation. Thus, knowledge of the phase diagram of this protein gave us a good indication of the conditions under which crystallization might occur (i.e., at higher temperatures).

QLS. We also carried out QLS measurements that showed that in the soluble region of the phase diagram, the diffusion coefficient of P23V is consistent with monomeric protein, and no aggregation is observed (region A in Fig. 1).

QLS and Protein Aggregation. In region B of the phase diagram (Fig. 1), where aggregation occurs, this aggregation was in the form of isolated particles/aggregates of protein, with larger hydrodynamic radii passing through the scattering volume. These large particles always represented a very small portion of

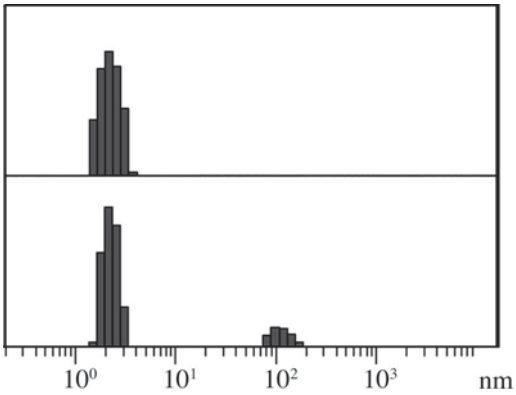

Fig. 3. Quasielastic light scattering for P23V. (*Upper*) P23V solution, 5 mg/ml, 42°C, 5 min after the beginning of measurement. Clearly the protein is monomeric. (*Lower*) P23V solution, 5 mg/ml, 42°C, 6 h after the beginning of measurement. Note the appearance of particles with a higher hydrodynamic radius, R_h, indicating aggregation of the protein. Even with protein aggregates present, a significant proportion of monomer remains in solution.

the entire sample, comprising usually <5% of the total scattering intensity (in the concentration and temperature regimes we examined). Most of the protein remained as monomer. It is important to note that there was no evidence of percolation or gelation in the light scattering data. In Fig. 3 we show QLS data for P23V at 5 mg/ml ($\phi = 0.0035$) at 42°C (315 K), just inside region B of the phase diagram. At the beginning of the measurement, only monomer is present in solution (Fig. 3 *Upper*). After a lag time, protein aggregates begin to form, and higher molecular weight aggregates are observed (Fig. 3 *Lower*). This aggregation is reversible. Once the temperature is lowered again, the aggregates dissolve, leaving only protein monomer in solution.

QLS in Region A of the Phase Diagram. We carried out QLS experiments in the soluble region of the phase diagram (Fig. 1, region A) at a number of different temperatures, to examine the "averaged" interaction between the proteins in solution. The results are shown in Fig. 4. We measured the collective diffusion coefficient (D_c) for P23V solutions between 4 and 30 mg/ml ($\phi =$ 0.0028 to 0.021) and at temperatures ranging from 20° to 35°C (293–308 K), all in region A of the phase diagram. For this range

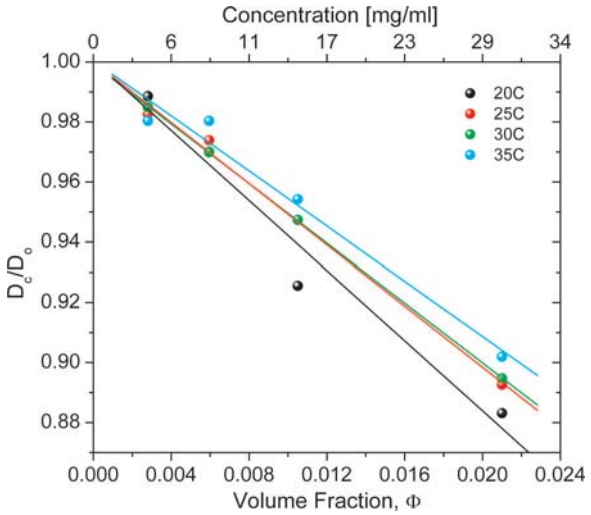

Fig. 4. $D_c(T, \phi)/D_0(T)$ vs. protein concentration (or volume fraction, ϕ) for P23V solutions at concentrations from 4 to 30 mg/ml at four temperatures.

Table 1. D_0 and interaction parameter, k_D, data shown as a function of temperature, calculated from a fit $D_c = D_0(1 + k_D\phi)$ to the data in Fig. 4

Temperature, K	D_0, cm²/s	k_D
293	0.97×10^{-6}	−5.83
298	1.15×10^{-6}	−5.09
303	1.33×10^{-6}	−4.97
308	1.53×10^{-6}	−4.59

of concentrations, we would expect D_c to depend linearly on ϕ. We linearly extrapolated D_c to $\phi = 0$ to determine the free-particle diffusion coefficient, D_0, which we found to be 1.15×10^{-6}cm²·s^{-1} (25°C, 298 K). This corresponds to a hydrodynamic radius of ≈2.1 nm by the Stokes–Einstein equation: $D_0 = kT/6\pi\eta R$, where D_0 is the diffusion coefficient (in cm²/s), η is the solvent viscosity, k is Boltzman's constant, and R is the hydrodynamic radius. These samples were all measured in the region of the phase diagram where aggregation and crystallization are absent and multiple scattering is not present. Our D_0 for HGD is consistent with D_0 for bovine γB-crystallin (23). There is a high degree of sequence homology between the γ-crystallins and, because the molecular weights of bovine γB-crystallin and P23V differ by only ≈350 Da, we would expect to measure very similar diffusion coefficients. Bovine γB-crystallin, like HGD, exhibits LLPS and has "normal" solubility.

QLS Measurements and Protein Interactions. Fig. 4 shows $D_c(\phi, T)/D_0(T)$ data for P23V at different temperatures. Our data show that D_c decreases with increasing protein concentration at all temperatures between 20° and 35°C (293–308 K) investigated. This behavior is expected from particles with attractive interactions because concentration fluctuations dissipate more slowly as the particle concentration increases, resulting in a smaller measured D_0. Our measurements of diffusivity for P23V show that with increasing temperature, the collective diffusion coefficient for the protein particles becomes smaller (at constant ϕ). Both the concentration and temperature dependence we find were observed by Fine *et al.* (23) for bovine γB-crystallin. In both cases, the slope of D_c vs. ϕ is negative and of approximately the same magnitude (approximately −7.5 for bovine γB-crystallin and approximately −5.0 for HGD), and the values for D_0 of both proteins at 25°C (298 K) agree within 10%. Several studies have used QLS to probe interactions between proteins (23–33). According to Muschol and Rosenberger (27),

$$D_c = D_0[1 + (k_S - k_H)\phi] \equiv D_0[1 + k_D\phi], \qquad [1]$$

where the static structure factor $S(q = 0, \phi) \approx 1 - k_S\phi$ and the hydrodynamic interaction term $H(q = 0, \phi) \approx 1 - k_H\phi$. k_S, k_H, and k_D are defined in ref. 27. We can deduce k_D at each temperature from our data. These results are shown in Table 1. We also include D_0 at each temperature. Examining the data, k_D is negative at all temperatures. We can compare this result with the value predicted for hard spheres, $k_D = 1.45$ [Batchelor (34)]. Clearly, our k_D has the opposite sign. As the temperature increases, our value for k_D becomes slightly less negative, indicating that the interactions between the protein particles become slightly less attractive (i.e., they are approaching a more hard-sphere, repulsive-type behavior). This solution behavior was not initially expected because of the strong inverted temperature dependence of the solubility. Both QLS data and LLPS data for native and mutant proteins indicate that the solution behavior is essentially the same for both proteins.

Comparison with HbS. We stated previously (1) that the retrograde solubility profile of the P23V mutant has also been observed in

the case of sickle cell hemoglobin, HbS (17). However, there is a clear distinction between the two cases: the nature of the solubility curve in HbS is neither a liquidus line nor a protein monomer–aggregate equilibrium as in P23V. The HbS solubility curve represents an equilibrium between protein monomers and protein fibers that have polymerized and exist in equilibrium with surrounding monomer. It is these polymer fibers that are responsible for the deformation of red blood cells in sickle cell anemia (17, 35). HbS also shows LLPS behavior, but it additionally has a lower consolute temperature. The features of the HbS phase diagram are, therefore, distinct from the phase behavior of P23V that we describe here.

Analysis of the Solubility Data. We can represent the data for the liquidus line for native HGD and the solubility line of the mutant proteins in terms of the magnitude and temperature dependence of the relevant transfer chemical potential, $\Delta\mu_{\rm trans}$, defined as

$$\Delta\mu_{\rm trans} = (\mu_P^0 + \kappa\mu_W^0) - \mu^{\rm C}(T, n_{\rm C}, \kappa), \qquad [2]$$

where $\mu^{\rm C}$ is the chemical potential of a protein molecule with κ associated water molecules in the solid phase. $\mu^{\rm C}$ generally represents the chemical potential in the crystal phase, but in this case we use it to represent the aggregate phase (for the reasons discussed earlier in this section). $\mu^{\rm C}$ is a negative number, larger in magnitude than the negative quantity, $\mu_{\rm P}^0 + \mu_{\rm W}^0$, which is the standard part of the chemical potential of the protein with κ associated water molecules. Therefore, $\Delta\mu_{\rm trans}$ is a positive quantity. We may use the Van't Hoff law, which states that the volume fraction, ϕ, along the solubility line is related to the transfer chemical potential, $\Delta\mu_{\rm trans}$, according to (36)

$$\ln\phi = -\left(\frac{\Delta\mu_{\rm trans}}{kT}\right). \qquad [3]$$

If we define $X = T_0/T$, where $T_0 = 298$ K, it follows that $\phi = \exp -(\Delta\mu_{\rm trans}/kT_0)X$. Using the solubility data for native HGD (6) and the mutants, P23V, P23S, and P23T (1), we can determine $\Delta\mu_{\rm trans}$ for each protein. Plotting $\Delta\mu_{\rm trans}/kT_0$ vs. T/T_0, we find that the slope for the native protein is negative, whereas it is positive for all of the mutant proteins. Also, using Eq. **3**, we observe that at volume fractions higher than $\approx$0.02, $\Delta\mu_{\rm trans}/kT_0$ does not depend linearly on T. It is important to determine the reason for the nonlinearity. Here we employ a more accurate form of Eq. **3**, using the virial expansion:

$$\ln\phi + 2B_2\phi = -\left(\frac{\Delta\mu_{\rm trans}}{kT_0}\right)X. \qquad [4]$$

B_2, the second virial coefficient, is the second term in the virial expansion. B_2 is defined by Lomakin *et al.* (37):

$$B_2 = -4[(e^{\varepsilon} - 1)(\lambda^3 - 1) - 1], \qquad [5a]$$

where $\varepsilon = \varepsilon_{\rm eff}/kT$, and from the LLPS, $\lambda = 1.25$ and $\varepsilon_{\rm eff}/kT_{\rm c} = 1.27$ (38). B_2 then becomes

$$B_2 = -4\left[0.953\left(e^{1.27\frac{T_c}{T}} - 1\right) - 1\right]. \qquad [5b]$$

We note that B_2 has a weak temperature dependence. Using this expression for B_2 in Eq. **4**, we find that $\Delta\mu_{\rm trans}/kT_0$ vs. T/T_0 is in fact linear over the temperature range studied (See Fig. 5). Defining $\tau = T/T_0$, we express this linear temperature dependence as

$$\frac{\Delta\mu_{\rm trans}(\tau)}{kT_0} = V + S\Delta\tau, \qquad [6]$$

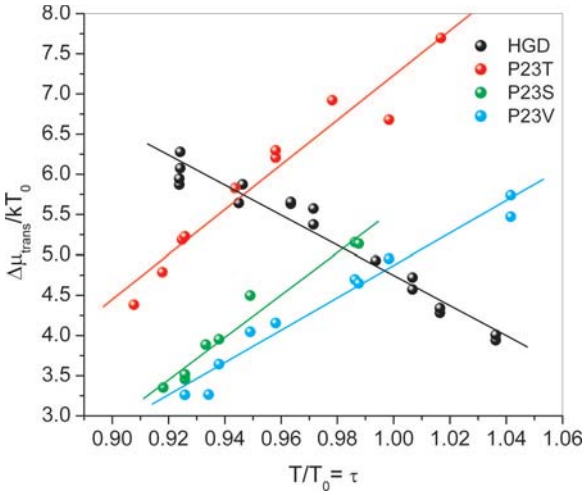

Fig. 5. Temperature dependence of $\Delta\mu^{\rm C}_{\rm trans}(T)/kT_0$ vs. T/T_0 for native HGD and the three mutant proteins, P23T, P23S, and P23V, as deduced by using Eq. **4**.

where $V = \Delta\mu_{\rm trans}(T_0)/kT_0$; $S = [\partial(\Delta\mu_{\rm trans}/kT_0)/\partial\tau]_{\tau=1}$; and $\Delta\tau = (T - T_0)/T_0$.

In Table 2 we list the values of S and V for each of the mutants and for the native protein, as determined from the data shown in Fig. 5. If we now use this linear dependence of $\Delta\mu_{\rm trans}/kT_0$ in the Van't Hoff law, and noting that $(T_0/T) \cong (1 - \Delta\tau)$, we find that in the low ϕ region ($\phi < \approx$0.02), the solubility line has the very simple form

$$\phi \cong \phi_0\exp[(V - S)\Delta\tau], \qquad [7a]$$

where

$$\phi_0 = \exp -\left(\frac{\Delta\mu_{\rm trans}(T_0)}{kT_0}\right). \qquad [7b]$$

From Eq. **7a** and **7b**, we see at once that the increase or decrease of the solubility ϕ with increasing temperature is determined entirely by the sign of the quantity $K = ([\Delta\mu_{\rm trans}(T_0)/kT_0] - S)$. The value of this quantity is listed in column 4 of Table 2 for the native and mutant proteins. In each case, the magnitude of the slope S makes the dominant contribution to K. For the native protein, K is +23.8, corresponding to the observed "normal" increase in solubility with increasing temperature. On the other hand, for the P23V mutant, K is negative because of the large positive value of S for this protein. Thus, the strong increase of $\Delta\mu_{\rm trans}/kT_0$ with increasing temperature is responsible for the retrograde solubility line for the P23V mutant and the other mutants as well.

The location of the solubility line in the (ϕ, T) plane is determined by the prefactor ϕ_0 according to Eq. **7b**. Indeed, the

Table 2. Values of the parameters *V* and *S*, describing respectively the $T = T_0$ intercept and the slope of $\Delta\ \mu_{\rm trans}(\tau)/kT_0$ vs. $\tau = T/T_0$ for each protein

Protein	*S*	*V*	$K = V - S$
HGD	−19	4.8	23.8
P23T	28	7.2	−20.8
P23S	26	5.6	−20.4
P23V	20	4.8	−15.2

Both *V* and *S* were determined by using a least-square fit. These values are reported to two significant figures in view of the scatter in the experimental data as shown in Fig. 5.

dramatic decrease in the observed solubility at T_0 as one compares P23V with P23S and P23T is, in fact, the result of an increase in V, from 4.8 to 7.2. Thus, the large change in solubility and consequent opacification of the lens is due to a relatively small change of approximately $\Delta\mu_{trans}(1) \approx 2.4kT_0$, which appears as the exponent in Eq. **7b**.

Aeolotopic Model and P23V Mutation. The striking feature of our experimental data is the finding that although the mutations dramatically change the solubility of the protein, they have a negligible effect on the properties of the solution phase, such as the LLPS or the concentration dependence of the protein diffusivity. In effect, the mutations change the chemical potentials of the solid phase but not those of the liquid phase. This represents an inexplicable paradox if the interprotein interactions are treated using a simple isotropic model. Indeed, in such a model a change in the attractive energy as a result of the mutation must have a proportional effect on both the coexistence curve and the solubility line. In reality, however, as we have emphasized previously (37), the interprotein interactions are highly anisotropic and localized. The analysis of our data indicates that a mutation at site 23 dramatically affects the temperature dependence of the interprotein interaction. Indeed, according to the estimates above, we deduce that the energy change upon mutation in this region has a strong temperature dependence, amounting to several kT over a ≈30°C temperature range. This energy magnitude is small compared with the typical energy associated with crystal contacts. Interestingly, the x-ray structure of native HGD crystals (39) indicates that site 23 is not directly involved as a crystal contact. Furthermore, no crystal structure is available for the site 23 mutants. Because the solubility at T_0 is nearly the same for mutant and native proteins, it is reasonable to assume that there is no significant structural change as a result of the mutation.

Our data show that a local energy change of this magnitude produces a negligible effect on the average pairwise interaction (second virial coefficient and diffusivity) and on the liquid–liquid coexistence in the solution phase. This observation finds its natural explanation in the framework of the aeolotopic model (37). In this model, we have shown that the thermodynamic properties of the liquid phase can be described by a spatial thermodynamic average of the highly anisotropic actual potential. Using the expression for the effective attractive energy given in ref. 37, suitably generalized for attractive spots with different energies and areas, it is possible to show that for a protein with >30 attractive spots, the aforementioned change of ≈2.4 kT_0 in the energy of one spot will translate into a <0.3% change in the thermodynamically averaged effective energy in the solution phase. This corresponds to ≈1°C shift in the coexistence curve, which is within the level of uncertainty in the measurements of LLPS. Thus, the apparent discrepancy between the interaction energies in the solution phase vs. the solid phase is explicable in terms of highly anisotropic interprotein interactions, which are averaged out in the solution phase but are fully engaged in the solid phase.

Role of Proline 23 in HGD. The overall structure of HGD has been determined (39), and Raman, IR, and far-UV CD analyses suggest that mutations at site 23 produce no major structural changes. Therefore, the observed dramatic changes in solubility, and especially the retrograde temperature dependence in the mutant proteins, must be ascribed to local changes in the immediate vicinity of site 23. The fact that, regardless of the residue replacing proline, the solubility curve always changes to retrograde, suggests the importance of the removal of the proline. Indeed, it has been found that the addition of proline at site 24, while the Val or Thr residues are left at position 23, restores the normal temperature dependence of the solubility of the protein (1). Thus we can conclude that the presence of proline in the vicinity of site 23 contributes importantly to the binding energy, decreasing it as the temperature increases. This conclusion presents an interesting challenge to theoretical analysis using molecular dynamic simulations.

Conclusions

We have reported the location of the boundary for LLPS and the solubility line, as well as the concentration and temperature dependence of the diffusion coefficient, in solutions of the P23V mutant of HGD. We have compared these properties with those of the native HGD and have found conditions under which crystals of the mutant P23V can be formed. As is the case with two other mutants at this position, P23V shows a retrograde solubility line, i.e., the solubility decreases with increasing temperature, in dramatic contrast to that of the native protein. As a result, at body temperature the solubility of the mutant is much less than that of the native protein. On the other hand, solution properties such as the liquid–liquid phase boundary and the collective diffusion coefficient are essentially unchanged by this mutation. The apparent inconsistency between the effect of mutation in the solution and solid phases can be understood in the framework of the aeolotopic model of highly anisotropic interactions between proteins. Analysis of the solubility line shows that the mutation causes changes in the binding energy in a strongly temperature-dependent fashion. As the temperature changes from 5° to 35°C, the binding energy of the mutant increases by ≈2.4 kT. Comparison with data on other mutants (1) suggests that the underlying source of these binding energy changes is associated with the presence or absence of proline in the vicinity of position 23.

Materials and Methods

Cloning, Expression, and Isolation of Proteins. Recombinant HGD was prepared by amplification of the coding sequence from a human fetal lens cDNA library, as described previously (1). Mutagenesis was performed with the QuikChange site-directed mutagenesis kit from Stratagene (La Jolla, CA), with primers synthesized by MWG Biotech (High Point, NC) also described elsewhere (1). The plasmid DNA obtained after mutagenesis was sequenced with the T7 promoter primer and found to contain the desired mutation and no other sequence changes. Overexpression of the recombinant proteins (HGD, P23V), and their isolation and purification, were carried out as described previously (1). In all cases, the crystallins were obtained from the soluble fraction. Each batch of protein prepared was analyzed by using electrospray ionization mass spectrometry (Biopolymers Laboratory, Center for Cancer Research, Massachusetts Institute of Technology). The average masses for proteins determined in this study were based on three batches of each protein and were HGD: 20,610 ± 1 and P23V: 20,612 ± 1 (consistent with ref. 1). For HGD and the mutant proteins used in this study, an extinction coefficient of 41.4 $mM^{-1}\cdot cm^{-1}$ at 280 nm was used (40). The volume fraction of the protein is expressed as $\phi = c\cdot\upsilon_{sp}$, where c is the concentration in milligrams per milliliter and υ_{sp} is the partial specific volume, 7.1×10^{-4} ml/mg (18).

QLS. QLS was performed with an instrument built in-house, using a coherent 35-mW, 632.8-nm He-Ne laser (Coherent Radiation, Santa Clara, CA) and a PD2000DLS^{PLUS} 256-channel correlator (Precision Detectors, Bellingham, MA). The scattering angle in all experiments was 90°. The temperature was controlled by using an external circulating water bath, with the sample temperature monitored by a thermocouple probe positioned beside the sample vial during measurements. The measured correlation functions were analyzed by the Precision Deconvolve 4.4 software method (Precision Detectors), which determines the total scattered light intensity vs. the diffusion coefficient of the

sample. Protein solutions were prepared at the desired concentration in 0.1 M sodium phosphate buffer (20 mM DTT/0.02% sodium azide) and then filtered through 0.02-μm Anatop filters (Whatman, Clifton, NJ) before measurement. Cylindrical glass tubes (Kimble, Vineland, NJ) with an internal diameter of 0.4 cm were used for QLS measurements.

Measurement of LLPS. P23V solutions, filtered through 0.22-μm PDVF filters (Millipore, Billerica, MA) of known concentration in 0.4-cm-diameter glass tubes, were placed in a cell with transparent windows, connected to an external circulating water bath. The samples were cooled to just below the solubility line ($\approx$4°C) and allowed to equilibrate. There is a very narrow region, below the solubility line and above the coexistence curve, where the protein exists as a homogeneous monomeric solution. It is important to ensure that the protein is in monomeric form before the temperature is lowered to measure LLPS, so particular care was taken in this regard. The temperature inside the cell was measured by a thermocouple probe placed in an identical glass vial containing water, inside the same cooled cell. A 4-mW He-Ne laser was focused on the sample, and the transmitted light intensity was detected by a photomultiplier connected to a voltmeter. The initial light intensity value was recorded. The cell was then cooled, and the transmitted light intensity was monitored. When the light intensity fell to half of its initial value, the temperature at which this occurred was recorded. When the transmitted intensity dropped further, below 1%, the cell was heated and the temperature at which the transmitted intensity reached half the initial value again was recorded. The phase separation temperature, T_{ph}, was taken as the average of the clouding and clearing temperatures. Interactions between γ-crystallins are attractive and are also characterized by a high content of free cysteine residues, which leaves these proteins susceptible to disulfide-bond formation. Here we have avoided the formation of these covalently linked aggregates with the addition of DTT (41).

We thank Dr. Nicolette Lubsen (Nijmegen Centre for Molecular Life Sciences, Nijmegen, The Netherlands) for providing the original human CRGD clone for all of our work and Dr. Neer Asherie for helpful discussions. This work was supported by National Institutes of Health Grants EY05127 (to G.B.B.) and EY10535 (to J.P.).

1. Pande A, Anunziata O, Asherie N, Ogun O, Benedek GB, Pande J (2005) *Biochemistry* 44:2491–2500.
2. Evans P, Wyatt K, Wistow GJ, Bateman OA, Wallace B, Slingsby C (2004) *J Mol Biol* 343:435–444.
3. Hejtmancik JF (1998) *Am J Hum Genet* 62:520–525.
4. Santhyiya ST, Shyam Manohar M, Rawlley D, Vijayalakshmi P, Namperumalsamy P, Gopinath PM, Loster J, Graw J (2002) *J Med Genet* 39:352–358.
5. Stephan DA, Gillanders E, Vanderveen D, Freas-Lutz D, Wisow G, Baxevanis AS, Robbins CM, VanAuken A, Wuesenberry MI, Bailey-Wilson J, *et al.* (1999) *Proc Natl Acad Sci USA* 96:1008–1012.
6. Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King JA, Lubsen NH, Walton D, Benedek GB (2000) *Proc Natl Acad Sci USA* 97:1993–1998.
7. Héon E, Priston M, Schorderet DF, Billingsley GD, Girard PO, Lubsen N, Munier FL (1999) *Am J Hum Genet* 65:1261–1267.
8. Kmoch S, Brynda J, Befekadu A, Bezouška K, Novák P, Rezácová P, Ondrová L, Filipec M, Sedlácek J, Elleder M (2000) *Hum Mol Genet* 9:1779–1786.
9. Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King J, Benedek GB (2001) *Proc Natl Acad Sci USA* 98:6116–6120.
10. Benedek G (1997) *Invest Ophthalmol Visual Sci* 38:1911–1921.
11. Ingram VM (1958) *Biochim Biophys Acta* 28:539–545.
12. Ingram VM (1959) *Biochim Biophys Acta* 36:402–411.
13. Lomakin A, Teplow DB, Kirschner DL, Benedek GB (1997) *Proc Natl Acad Sci USA* 94:7942–7947.
14. Walsh DM, Lomakin A, Benedek GB, Condron MM, Teplow DB (1997) *J Biol Chem* 272:22364–22372.
15. Hull RL, Westermark GT, Westermark P, Kahn SE (2004) *J Clin Endocrinol Metab* 89:3629–3643.
16. Skvoronsky DM, Lee VM-Y, Trojanowski JQ (2006) *Annu Rev Pathol Mech Dis* 1:151–170.
17. Serrano MD, Galking O, Yaua S-T, Thomasa BR, Nagel RL, Hirsche RE, Vekilov PG (2001) *J Cryst Growth* 232:368–375.
18. Broide ML, Berland CR, Pande J, Ogun O, Benedek GB (1991) *Proc Natl Acad Sci USA* 88:5660–5664.
19. Berland CR, Thurston GM, Kondo M, Broide ML, Pande J, Ogun O, Benedek GB (1992) *Proc Natl Acad Sci USA* 89:1214–2849.
20. Asherie N (2004) *Methods* 34:266–272.
21. Feeling-Taylor RA, Banish RM, Hirsch RE, Vekilov PG (1999) *Rev Sci Instrum* 70:2845–2849.
22. Boistelle R, Astier JP, Marchis-Mouren G, Besseaux V, Haser R (1992) *J Cryst Growth* 123:109–120.
23. Fine BM, Lomakin A, Ogun O, Benedek GB (1996) *J Chem Phys* 104:326–335.
24. Piazza R (2000) *Curr Opin Colloid Interface Sci* 5:38–43.
25. Da Xing SL, Li J (2004) *J Biol Phys* 30:313–324.
26. Placidi M, Cannistraro S (1998) *Europhys Lett* 43:476–481.
27. Muschol M, Rosenberger F (1995) *J Chem Phys* 103:10425–10432.
28. Zhang J, Liu XY (2003) *J Chem Phys* 119:10972–10976.
29. Pellicane G, Costa D, Caccamo C (2004) *J Phys Chem B* 108:7538–7541.
30. Kanzaki N, Uyeda TQP, Onuma K (2006) *J Phys Chem B* 110:2881–2887.
31. Pusey PN (1993) in *Dynamic Light Scattering, The Method and Some Applications*, ed Brown W (Oxford Univ Press, New York), p 85.
32. Goldstein DL (1975) *States of Matter* (Prentice–Hall, Englewood Cliffs, NJ), p 265.
33. Magid L (1993) in *Dynamic Light Scattering, The Method and Some Applications*, ed Brown W (Oxford Univ Press, New York), p 554.
34. Batchelor JK (1976) *J Fluid Mech* 74:1–29.
35. Galkin O, Chen K, Nagel RL, Elison Hirsch R, Vekilov PG (2002) *Proc Natl Acad Sci USA* 99:8479–8483.
36. McQuarrie DA (2000) in *Statistical Mechanics* (University Science Books, Sausalito, CA), p 151.
37. Lomakin A, Asherie N, Benedek GB (1999) *Proc Natl Acad Sci USA* 96:9465–9468.
38. Lomakin A, Asherie N, Benedek GB (2003) *Proc Natl Acad Sci USA* 100:10254–10257.
39. Basak A, Bateman O, Slingsby C, Pande A, Asherie N, Ogun O, Benedek GB, Pande J (2003) *J Mol Biol* 328:1137–1147.
40. Andley UP, Mathur S, Griest TA, Petrash JM (1996) *J Biol Chem* 271:31973–31980.
41. Pande J, Lomakin A, Fine B, Ogun O, Sokolinski I, Benedek GB (1995) *Proc Natl Acad Sci USA* 92:1067–1071.