Effect of polyethylene glycol on the liquid–liquid phase transition in aqueous protein solutions

Onofrio Annunziata, Neer Asherie, Aleksey Lomakin, Jayanti Pande, Olutayo Ogun, and George B. Benedek*

Department of Physics, Center for Materials Science and Engineering, and Materials Processing Center, Massachusetts Institute of Technology, Cambridge, MA 02139-4307

Contributed by George B. Benedek, August 22, 2002

We have studied the effect of polyethylene glycol (PEG) on the liquid–liquid phase separation (LLPS) of aqueous solutions of bovine γD-crystallin (γD), a protein in the eye lens. We observe that the phase separation temperature increases with both PEG concentration and PEG molecular weight. PEG partitioning, which is the difference between the PEG concentration in the two coexisting phases, has been measured experimentally and observed to increase with PEG molecular weight. The measurements of both LLPS temperature and PEG partitioning in the ternary γD-PEG-water systems are used to successfully predict the location of the liquid–liquid phase boundary of the binary γD-water system. We show that our LLPS measurements can be also used to estimate the protein solubility as a function of the concentration of crystallizing agents. Moreover, the slope of the tie-lines and the dependence of LLPS temperature on polymer concentration provide a powerful and sensitive check of the validity of excluded volume models. Finally, we show that the increase of the LLPS temperature with PEG concentration is due to attractive protein–protein interactions.

PEG | ternary mixtures | solubility | partitioning

Polyethylene glycol (PEG) is a hydrophilic nonionic polymer used in many biochemical and industrial applications. Due to its nontoxic character, this chemical can be found in cosmetics, food, and pharmaceutical products. The mild action of PEG on the biological activity of cell components explains the success of this polymer in biotechnological applications. PEG is commonly used for liquid–liquid partitioning and precipitation of biomacromolecules (1, 2). In protein crystallography, PEG is considered the most successful precipitating agent for the production of protein crystals, the crucial step for the determination of the molecular structure of a protein. All these applications make PEG by far the most widely used polymer in aqueous solutions of biological molecules (1, 3).

Due to the extensive practical use of PEG as a precipitating agent for proteins, it is of fundamental importance to understand protein–protein and protein–PEG interactions in protein–PEG–water ternary systems. These interactions are often described in terms of a depletion force, which arises because the polymer is depleted in the region between adjacent proteins (4, 5). Depletion force models have been successful in describing the effect of nonadsorbing polymers on colloidal suspensions (6–10).

Protein–PEG–water solutions have been investigated by several techniques (4, 5, 11–14). Small-angle x-ray- (4) and light-scattering (5) measurements generally confirmed that total protein–protein interactions can be described in terms of depletion effects. The microscopic interpretation used in the above studies is derived from colloid–polymer models. To model the effect of PEG, a depletion component is added to the original protein–protein pair interaction potential. This microscopic modeling, however, does not provide information about the actual protein–PEG interactions in the ternary system. Dialysis experiments, which describe the effect of PEG partitioning (i.e., the difference of PEG concentration) between two solutions separated by a membrane not permeable to protein molecules, give information on protein–PEG interactions only when protein–protein interactions are absent (12, 15). Studies of the solubility of protein crystals in the presence of PEG have been carried out to investigate protein–protein and protein–PEG interactions. However, to interpret the results, assumptions about the properties of the crystal are required. Moreover, simple excluded volume models fail to describe protein solubility as a function of PEG molecular weight (16). The discrepancy between model and experiment becomes particularly large for PEG of lower molecular weights (11). Furthermore, the magnitude of the protein–PEG interactions changes from protein to protein in a way that does not correlate with the simple excluded volume scenario (11, 15).

Liquid–liquid phase separation (LLPS) can be used as a direct tool for probing both protein–protein and protein–PEG interactions. It has been shown that the position of the liquid–liquid boundaries provides insight about the nature and magnitude of the protein–protein interactions (17–20). The advantage of LLPS analysis with respect to measurements of protein solubility is that data interpretation requires no assumptions regarding crystal phase properties (11, 21). LLPS is important also for the analysis of nonequilibrium properties of the system. Because the nucleation of protein crystals is enhanced in proximity of the LLPS, this phase boundary is believed to be implicated in protein crystallization kinetics (22, 23).

We present an experimental investigation of the effect of low molecular weight PEG on the LLPS of aqueous solutions of bovine γD-crystallin (γD). γD is a member of the γ-crystallin family of lens proteins, which are involved in cataract formation (24). The phase diagram of aqueous γ-crystallins has been well characterized, and several members of the family, including γD, have attractive interactions leading to LLPS of the corresponding aqueous solutions (17, 20, 25, 26). γD is a protein with a simple globular structure (27) and is weakly charged at the pH value (≈7) used in our measurements (28). The experiments were performed at a PEG concentration of ≈5% in weight, which falls within the range of concentrations used for protein crystal growth (3).

The LLPS properties of the protein–PEG–water ternary system are described by a coexistence surface in the phase diagram. This coexistence surface represents the LLPS temperature as a function of protein concentration, c_1, and PEG concentration, c_2. At a fixed temperature, the LLPS properties of the system are described by an isothermal coexistence curve, which gives the concentrations $(c_1^{\mathrm{I}}, c_2^{\mathrm{I}})$ and $(c_1^{\mathrm{II}}, c_2^{\mathrm{II}})$ of the coexisting phases, I and II. The partitioning of the components in the two coexisting phases is described by tie-lines connecting the points $(c_1^{\mathrm{I}}, c_2^{\mathrm{I}})$ and $(c_1^{\mathrm{II}}, c_2^{\mathrm{II}})$ of the coexistence curve. The critical point, $(c_1^{\mathrm{c}}, c_2^{\mathrm{c}})$, is defined as the point of the coexistence curve where the condition $(c_1^{\mathrm{I}}, c_2^{\mathrm{I}}) = (c_1^{\mathrm{II}}, c_2^{\mathrm{II}})$ occurs. The location of this point as a function of the temperature is described by a critical line on the coexistence surface.

Abbreviations: PEG, polyethylene glycol; LLPS, liquid–liquid phase separation; γD, γD-crystallin.

*To whom correspondence should be addressed. E-mail: gbb@mit.edu.

BIOPHYSICS

Materials and Methods

Bovine γD, which has a molecular mass of 20,700 g/mol, was isolated from 1- to 6-week-old calf lenses, obtained by overnight express from Antech (Tyler, TX). Pure γD was isolated and purified from these lenses by standard procedures described elsewhere (25, 29). The purity of the native sample was at least 95%, based on both ion-exchange and size-exclusion HPLC. The purified γD was dialyzed exhaustively into sodium phosphate buffer (0.1 M, pH 7.1) that contained sodium azide (0.02%).

PEG with molecular masses ranging from 200 to 3,350 g/mol (PEG200–PEG3350) and tetraethylene glycol (TEG, molecular mass 194.2 g/mol) were purchased from Sigma–Aldrich and used without further purification. Sample molecular mass and polydispersity were analyzed by electrospray ionization mass spectrometry at the Biopolymers Laboratory at the Center for Cancer Research at the Massachusetts Institute of Technology. Estimates of the average molecular mass were in agreement with the nominal values provided by the company. All polymers had a narrow molecular mass distribution with a polydispersity of about 1.05.

Protein–PEG aqueous solutions were prepared as follows. Solutions containing dilute γD in phosphate buffer were concentrated by ultrafiltration. When the desired protein concentration was reached, a known weight of PEG was added to the protein solution. The concentration of γD in the samples was determined by UV absorption at 280 nm, using the extinction coefficient value of 2.11 $mg^{-1}{\cdot}ml{\cdot}cm^{-1}$ (26). The concentration of PEG in the samples was calculated by using the mass of PEG and the total volume of the solution. The volumetric contribution of each component is obtained by multiplying the mass of the component by the corresponding specific volume, i.e., 0.71 ml/g for γD (30), 0.84 ml/g for PEG (15) and 0.992 ml/g for the buffer.

The LLPS temperature for a given protein–PEG aqueous solution was determined by gradually lowering the temperature of the sample until clouding was observed. The cloud point was determined by the examination of the transmitted intensity–temperature profile as previously described by Liu *et al.* (31). The coexisting phases were obtained by quenching the sample at fixed temperature beneath the coexistence surface. If after about 24 hours the two coexisting phases had not separated by gravity, centrifugation was used for the separation. The protein concentration in each phase was determined by UV absorption. To determine the PEG concentration in each of the coexisting phases, an aliquot of known weight was taken from each phase, and the PEG was separated from the protein by ultrafiltration (Amicon Microcon YM-10). The protein-free solution was weighed and the corresponding concentration of PEG determined by using a standardized refractive index detector (Perkin–Elmer LC-30 RI). The concentration of PEG in each aliquot was then calculated from its measured value in the protein-free solution. To ensure that possible differences in phosphate concentration in the two phases did not affect the results, the refractive index measurements were performed after isocratic elution of the protein-free solution on a size-exclusion HPLC column, by using sodium phosphate buffer (0.1 M, pH 7.1) with a flow rate of 1 ml/min. An Ohpak SB-802.5 HQ column (size 8 mm × 300 ml) from Phenomenex was used for the PEG200 and PEG400 solutions. For PEG1000, the Superdex 75 HR 10/30 column (size 10 × 300 mm) from Amersham Pharmacia Biotech was preferred for its higher resolution. The procedure was verified with protein–PEG aqueous solutions of known composition. In all cases, the measured protein and PEG concentrations in the two coexisting phases were consistent with the protein and PEG concentrations in the original homogeneous samples.

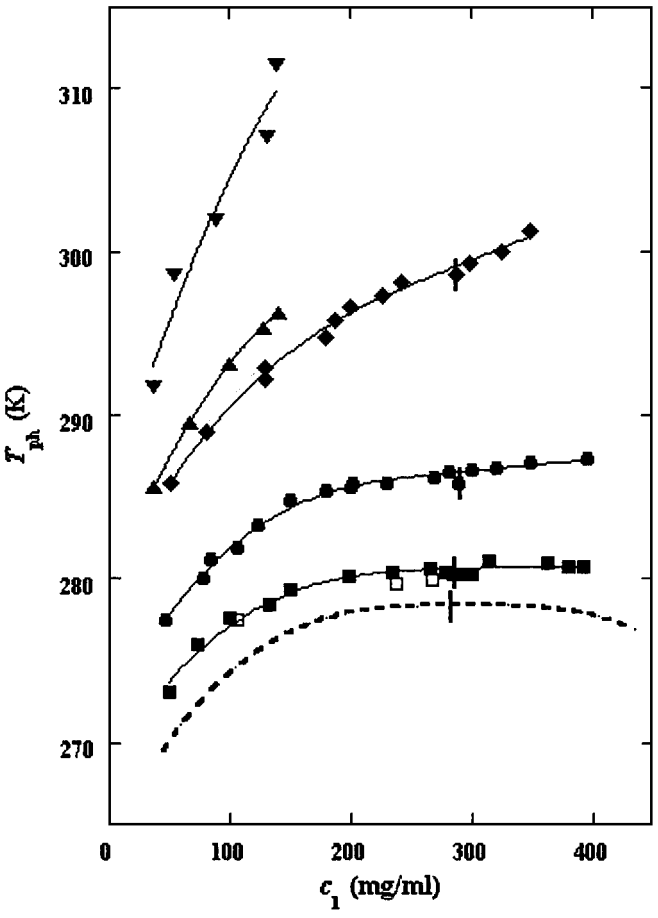

Fig. 1. LLPS temperature at constant PEG concentration of ≈50 mg/ml for the γD-PEG-water ternary system. The average molecular masses are: 200 g/mol (■), 400 g/mol (●), 1,000 g/mol (◆), 1,450 g/mol (▲), and 3,350 g/mol (▼). The solid curves are guides for the eye. The values for the γD-tetraethylene glycol-water ternary system are represented by open squares (□). We draw the coexistence curve for the γD-water binary system (dashed curve). The vertical bars (|) locate the critical point.

Results

We describe the composition of the γD-PEG-water systems by the protein (component 1) concentration c_1 and the PEG (component 2) concentration inside the protein-free volume, c_{2s}. If ν is the protein specific volume and $\phi = c_1\nu$ is the protein volume fraction, then c_{2s} is equal to $c_2/(1 - \phi)$. The choice of c_{2s} instead of c_2 to present our results is justified by the following argument. We consider as a reference system one in which the particles of component 2 have a negligible size and do not introduce specific chemical interactions in the system. In this reference case, the different values of c_2 between the two coexisting phases are entirely determined by the exclusion of component 2 from the volume occupied by component 1. Here, the condition of chemical equilibrium is expressed by the relation: $c_{2s}^{I} = c_{2s}^{II}$. Thus, in our systems, any observed difference in c_{2s} between two coexisting phases is related to the finite size and specific chemical properties of the PEG molecules.

In Fig. 1, we present our measurements of the LLPS temperature, T_{ph}, for γD-PEG-water ternary solutions at constant PEG concentration of $c_{2s} \approx 50$ mg/ml and with several polymer average molecular masses: 200 g/mol (filled squares), 400 g/mol (circles), 1,000 g/mol (diamonds), 1,450 g/mol (upright triangles), and 3,350 g/mol (inverted triangles). We also report measurements of temperature of phase separation for the γD-

Table 1. The values of the parameters characterizing the ternary coexistence surfaces

	c_{2s}, mg/ml	T_c, K	$(\partial T_{ph}/\partial c_1)^c_{c2s}$, 10^{-3} K mg^{-1}·ml	$(\partial c_{2s}/\partial c_1)^c_{T_{ph}}/c^c_{2s}$, 10^{-3}·mg^{-1}·ml	$(\partial T_{ph}/\partial c_{2s})^c_{c1}$, K^{-1}·mg^{-1}·ml	T^0_c, K
PEG200	54	280.6	3 ± 2	−0.5 ± 0.1	0.11 ± 0.07	275 ± 4
PEG400	52	286.5	9 ± 1	−1.1 ± 0.1	0.17 ± 0.03	278 ± 2
PEG1000	50	299.3	33 ± 3	−2.0 ± 0.3	0.33 ± 0.05	283 ± 3

tetraethylene glycol-water at the same value of c_{2s} (open squares). TEG essentially has the same molecular weight as PEG200, but, in contrast to PEG, is monodisperse. We see from the figure that the two sets of measurements, for TEG and for PEG200, overlap within the experimental error and conclude that the effects due to polydispersity are small. In Fig. 1, we also show the coexistence curve for the binary aqueous protein system (dashed curve) (29). The phase boundaries of the ternary systems are located at higher T_{ph} than that of the binary system, and the difference increases with PEG molecular weight. The critical points, shown by vertical bars in Fig. 1, were identified by the examination of the transmitted intensity–temperature profile as previously described by Liu *et al.* (31). In all cases, the value of protein critical concentration (about 290 mg/ml) is not affected by PEG within the experimental error. On the other hand, the slope of T_{ph} at the critical point, $(\partial T_{ph}/\partial c_1)^c_{c2s}$, is positive and increases with the PEG molecular weight, in contrast with the zero slope required for the protein–water binary system. In Table 1, we report the values of the PEG concentration, c_{2s}, the critical temperature, T_c, and the slope $(\partial T_{ph}/\partial c_1)^c_{c2s}$.

In Fig. 2, we present our experimental data (filled circles) of the coexisting values (c_1^I, c_{2s}^I) and (c_1^{II}, c_{2s}^{II}) for the γD-PEG-water solutions having PEG average molecular weight: 200 g/mol (279.2 K, case A), 400 g/mol (281.7 K, case B), and 1000 g/mol (290.3 K, case C). The pairs of points representing the coexisting phases are connected by straight lines (tie-lines). The dashed curves are eye guides representing the coexisting curves. In all three cases, the coexisting phases were obtained by quenching samples with the same initial protein concentration of 300 mg/ml but with different initial PEG concentrations. The temperatures for the measurements of the coexistence curves were chosen so that they lie above the critical temperature of the protein–water binary system but below the critical temperature of the corresponding ternary system in Fig. 1. Because the protein critical concentration, c_1^c, is equal to the average concentration $(c_1^I + c_1^{II})/2$ in the limit of $c_1^I - c_1^{II} = 0$, we determine c_1^c by plotting $(c_1^I + c_1^{II})/2$ as a function of $c_1^I - c_1^{II}$. The resulting values of the critical concentrations (vertical bars in Fig. 2) are, within the experimental error, the same as the value reported for the binary protein–water system. An interesting feature of the phase diagram is that the difference of PEG concentration between the two coexisting phases increases as the PEG molecular weight increases. To describe the effect of PEG molecular weight on PEG partitioning, we consider the normalized slope of the tie-lines at the critical point: $(\partial c_{2s}/\partial c_1)^c_{T_{ph}}/c^c_{2s}$. This quantity can be obtained by plotting the incremental ratio $(\ln c_{2s}^I - \ln c_{2s}^{II})/(c_1^I - c_1^{II})$ as a function of $c_1^I - c_1^{II}$ and considering the limit of $c_1^I - c_1^{II} = 0$. In Table 1, we can see that $(\partial c_{2s}/\partial c_1)^c_{T_{ph}}/c^c_{2s}$ increases with the PEG molecular weight. The protein volume fraction in the protein-rich phase is comparable to the protein volume fraction in the crystal, which is ≈0.4–0.6 ($c_1 \approx$ 600–800 mg/ml) (27, 32). The nonnegligible values of PEG concentration of the protein-rich phases reported for the γD-PEG1000-water system suggest that the crystal phase could accommodate polymer coils with molecular weight equal to or lower than ≈1,000 g/mol.

The behavior of the coexistence surface is described by the three slopes $(\partial c_{2s}/\partial c_1)_{T_{ph}}$, $(\partial T_{ph}/\partial c_{2s})_{c_1}$, and $(\partial T_{ph}/\partial c_1)_{c_{2s}}$, which are related to each other by the mathematical relationship:

$$\left(\frac{\partial T_{ph}}{\partial c_1}\right)_{c_{2s}} = -\left(\frac{\partial T_{ph}}{\partial c_{2s}}\right)_{c_1}\left(\frac{\partial c_{2s}}{\partial c_1}\right)_{T_{ph}}. \quad [1]$$

It is important to observe that Eq. **1** is true only for ternary systems. Due to the presence of the two buffer salts (dibasic and

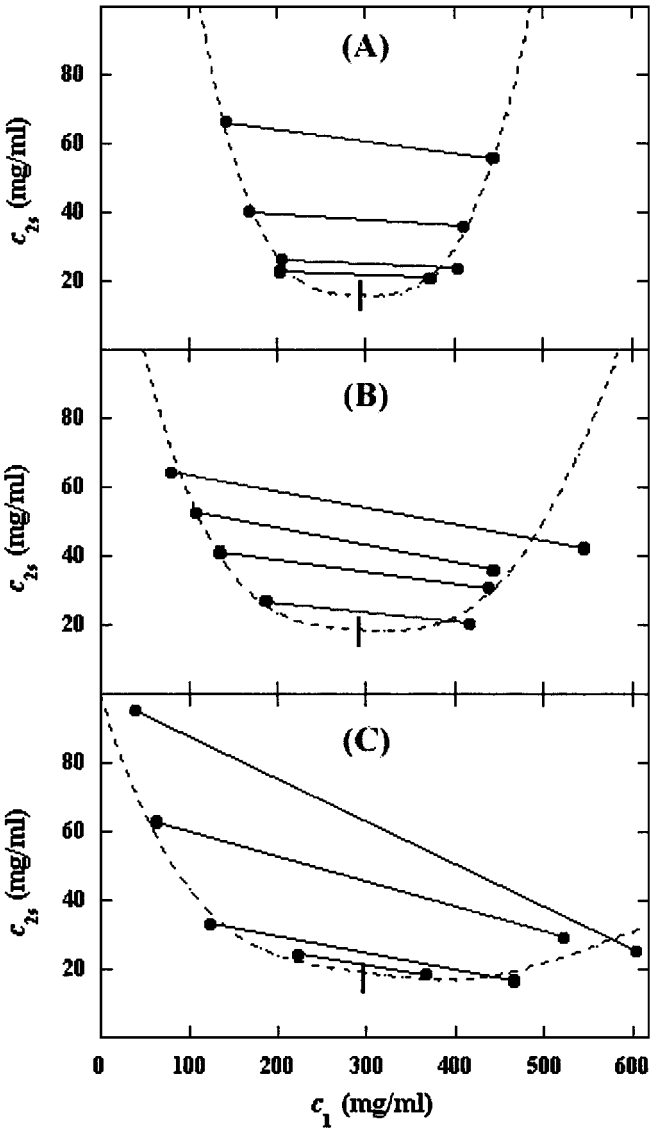

Fig. 2. Coexisting surfaces at constant temperature for the γD-PEG-water ternary system with PEG average molecular mass: 200 g/mol (279.2 K, case A), 400 g/mol (281.7 K, case B), and 1,000 g/mol (290.3 K, case C). The pairs of points representing the coexisting phases (●) are connected by the tie-lines (solid lines). The dashed curves are guides for the eye, and the vertical bars (|) locate the critical point.

monobasic sodium phosphate), our protein aqueous solutions contain more than three components. We know that salt partitioning is unimportant if PEG is not in the solution, because the protein-buffer system is well described as a binary system [for example, the critical point is at the maximum of the LLPS curve (26)]. Therefore, in the protein–PEG–buffer system, it is only because of the presence of PEG that salt partitioning may occur. We will show that Eq. **1** successfully describes our experimental results, which suggests that salt partitioning is insignificant insofar as the validity of Eq. **1** is concerned.

We apply Eq. **1** on the critical line of the coexistence surfaces. Because the critical temperature, T_c, of the γD–PEG–water systems (with $c_{2s} \approx 50$ mg/ml) is higher than the critical temperature, T_c^0, of the γD-water system (Fig. 1), the quantity $(\partial T_{ph}/\partial c_{2s})^c_{c_1}$ is positive. At the critical point, the slope $(\partial c_{2s}/\partial c_1)^c_{T_{ph}}$ of the coexistence surface is negative because it coincides with the slope of the tie-lines (Fig. 2). Eq. **1** shows that the slope $(\partial T_{ph}/\partial c_1)^c_{c_{2s}}$ is positive, as found experimentally (Fig. 1 and Table 1).

The slope $(\partial T_{ph}/\partial c_{2s})^c_{c_1}$ can be calculated, by using Eq. **1**, from the values of c_{2s}, $(\partial T_{ph}/\partial c_1)^c_{c_{2s}}$, and $(\partial c_{2s}/\partial c_1)^c_{T_{ph}}/c^c_{2s}$ reported in Table 1. We have assumed that the quantity, $(\partial c_{2s}/\partial c_1)^c_{T_{ph}}/c^c_{2s}$, which is obtained from our measurements of coexistence curves, is not a function of the temperature and PEG concentration within our experimental domain. The calculated values of $(\partial T_{ph}/\partial c_{2s})^c_{c_1}$, reported in Table 1, are consistent with the experimentally observed dependence of the critical temperature as a function of PEG molecular weight (Fig. 1). Indeed, we can quantitatively compare the experimental change of critical temperature between the ternary and the binary systems, shown in Fig. 1 for each PEG molecular weight, with the change predicted from the calculated values of $(\partial T_{ph}/\partial c_{2s})^c_{c_1}$. Specifically, we may estimate the critical temperature, T_c^0, of the γD-water system from the expression: $T_c^0 \approx T_c - (\partial T_{ph}/\partial c_{2s})^c_{c_1} c_{2s}$. Thereby, we obtain the calculated values of T_c^0, which are reported in last column of Table 1. These are in acceptable agreement with the experimental value of 278.4 K (29). Hence, using only the information obtained from the properties of the ternary coexistence surfaces, we are able to estimate reliably the location of the binary liquid–liquid phase boundary. This information can be most valuable when LLPS of protein–water systems is not experimentally observable because it is inaccessible due to either the freezing of the system or the denaturation of the protein at high temperature, which is important for proteins with lower consolute critical points (33).

Discussion

Our previous analysis shows that the behavior of the coexistence surfaces at the critical point can be described by two distinct features: the slope of the tie-lines and the dependence of LLPS temperature on PEG concentration. To analyze the physical factors that determine these two features of the coexistence surfaces, we introduce the free energy of the system. We define the quantity F, representing the difference, at constant volume, V, and temperature, T, between the Helmholtz free energy of the virtually incompressible protein–PEG–water system and the pure water system. The changes of F due to the replacement (at constant volume) of γ_1 water moles by 1 mole of protein and γ_2 water moles by 1 mole of PEG are respectively described by the differences of chemical potentials, $\mu_1 \equiv \mu_{protein} - \gamma_1\mu_{water}$ and $\mu_2 \equiv \mu_{PEG} - \gamma_2\mu_{water}$. A ternary incompressible system may be equivalently treated as a binary compressible system where μ_1 and μ_2 are the chemical potentials of the two effective components (17). The quantities μ_1 and μ_2 will be indicated as the protein and PEG effective chemical potentials.

It is convenient to introduce the reduced free energy $\hat{f} \equiv (F - F^0)/RTV$, where F^0 is the standard free energy, and R is the ideal gas constant. If the concentration of PEG, c_2, is relatively small, the reduced free energy is, to first-order approximation with respect to c_2, given by the following equation:

$$\hat{f}(c_1, c_2, T) = \hat{f}(c_1, 0, T) + c_2 \ln(c_2/e) + c_2\xi(c_1, T). \quad [2]$$

In Eq. **2**, $\hat{f}(c_1, 0, T)$ is the reduced free energy of the binary protein–water system, whereas $c_2 \ln(c_2/e)$ and $c_2\xi(c_1, T)$ are the contributions to the reduced free energy associated with the replacement of water molecules by PEG molecules at constant V and T. The contribution of PEG to the ideal mixing entropy of the system is represented by the term $c_2 \ln(c_2/e)$, whereas the quantity $c_2\xi(c_1, T)$ is the first term in a series expansion describing the change of the reduced free energy over and above the ideal mixing entropy. This approximation is acceptable because our experimental PEG concentrations are generally well below the threshold for polymer–polymer interpenetration, i.e., we are in the dilute regime (5). If we differentiate Eq. **2** with respect to the concentrations, c_1 and c_2, at constant V and T, we obtain the following expressions for the reduced effective chemical potentials:

$$\hat{\mu}_1(c_1, c_2, T) = \hat{\mu}'_1(c_1, T) + c_2(\partial\xi/\partial c_1)_T, \quad [3a]$$

$$\hat{\mu}_2(c_1, c_2, T) = \ln c_2 + \xi(c_1, T), \quad [3b]$$

where $\hat{\mu}_i \equiv (\mu_i - \mu_i^0)/RT$ (with $i = 1, 2$), $\mu_i^0(T) \equiv (\partial F^0/\partial c_i)_{T,V}/V$, and the quantity $\hat{\mu}'_1(c_1, T) \equiv \hat{\mu}_1(c_1, 0, T)$ is the protein effective chemical potential corresponding to the protein–water binary system. The two crossderivatives of the chemical potentials, $(\partial\hat{\mu}_1/\partial c_2)_{c_1,T}$ and $(\partial\hat{\mu}_2/\partial c_1)_{c_2,T}$, are both equal to $(\partial\xi/\partial c_1)_T$. This quantity characterizes the protein–PEG interactions. Eq. **3a** and **3b** fully characterize all the experimental observable features of the coexistence surfaces. Specifically, they enable us to analyze the quantities $(\partial c_2/\partial c_1)_{\hat{\mu}_2,T}$ (i.e., the slope of the tie-lines) and $(\partial T_{ph}/\partial c_2)_{c_1}$.

The Slope of the Tie-Lines. If we differentiate Eq. **3a** with respect to c_2 at constant c_1 and T, and Eq. **3b** with respect to c_1 at constant $\hat{\mu}_2$ and T, we obtain:

$$\left(\frac{\partial\hat{\mu}_1}{\partial c_2}\right)_{c_1,T} = \left(\frac{\partial\xi}{\partial c_1}\right)_T = -\frac{1}{c_2}\left(\frac{\partial c_2}{\partial c_1}\right)_{\hat{\mu}_2,T}. \quad [4]$$

Experimentally, the quantity $(\partial c_2/\partial c_1)_{\hat{\mu}_2,T}$ can be found from the limiting value of the slope of the tie-lines at the critical point. If we apply Eq. **4** to our experimental tie-line, we observe that $(\partial\hat{\mu}_1/\partial c_2)_{c_1,T} = (\partial\xi/\partial c_1)_T$ is positive and increases with the PEG molecular weight.

The slope of the tie-lines can be used to predict the effect of PEG on protein solubility. If protein crystals (solid phase) are in thermodynamic equilibrium with the liquid phase, $\hat{\mu}_1$ must be equal to its value in the solid phase, $\hat{\mu}_1^S$. For relatively low protein concentrations, $\hat{\mu}'_1(c_1, T) \approx \ln c_1$. Thus, using the equality of chemical potentials and Eq. **3a**, we find: $\ln c_1 \approx \hat{\mu}_1^S - (\partial\xi/\partial c_1)_T c_2$ (11). This relationship gives the protein solubility, c_1, as a function of PEG concentration, c_2. The dependence of $\hat{\mu}_1^S$ on c_1 and c_2 is expected to be very weak. Thus the fractional decrease of protein solubility with PEG concentration is fully determined by the value of $(\partial\xi/\partial c_1)_T$, which is proportional to the slope of the tie-line (see Eq. **4**). Therefore the magnitude of $(\partial c_2/\partial c_1)_{\hat{\mu}_2,T}/c_2$ characterizes the effectiveness of a protein precipitating agent such as PEG.

The Excluded Volume Model of Protein–PEG Interactions. We now analyze the experimental coexisting values (ϕ^I, c_2^I) and (ϕ^{II}, c_2^{II}) to gain insight about the physical factors that determine the quantity ξ and its dependence on c_1. We apply a simple excluded volume model that directly relates PEG partitioning to the

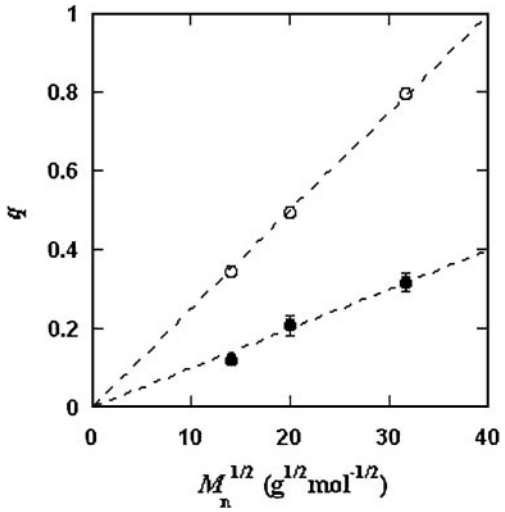

Fig. 3. Dependence of the size ratio, q, as a function of the square root of the PEG average molecular weight, M_n. The filled circles (●) are the value obtained by fitting the experimental coexisting compositions with Eq. **2**; the open circles (○) are the values calculated from the polymer gyration radii and the protein molecular volume. The dashed lines are guides for the eye.

difference in free volume fractions between the two coexisting phases. If we assume that PEG molecules can be described as ideal polymer coils and that protein molecules have a spherical shape, then due to steric hindrance, each protein will be surrounded by an adjacent region where the centers of mass of the coils are excluded, and the width, δ, of the resulting depletion layer will be proportional to the gyration radius, R_g, of the polymer coil (4–14). If α is the volume fraction available to the centers of mass of the coils, the condition of chemical equilibrium $\hat{\mu}_2(c_1^{\rm I}, c_2^{\rm I}, T) = \hat{\mu}_2(c_1^{\rm II}, c_2^{\rm II}, T)$ becomes $c_2^{\rm I}/\alpha^{\rm I} = c_2^{\rm II}/\alpha^{\rm II}$, where c_2/α is the polymer concentration in the free volume (9, 34) and the free volume fraction, α, is equal to $\exp(-\xi)$ (see Eq. **3b**).

It has been theoretically shown that, in the case of ideal polymer coils and relatively large hard spheres, $\delta \cong 1.1R_g$ (35). An approximate expression for the free volume fraction, α, as a function of the hard sphere volume fraction, ϕ, is given by the well established scaled particle theory (34, 36, 37):

$$\alpha = (1 - \phi)\exp(-A\eta - B\eta^2 - C\eta^3), \qquad [5]$$

where $\eta = \phi/(1 - \phi)$, $A = 3q + 3q^2 + q^3$, $B = 9q^2/2 + 3q^3$, $C = 3q^3$, and $q = \delta/R$ is the depletion layer thickness normalized with respect to the hard sphere radius, R. The prefactor, $1 - \phi$, in Eq. **5**, is the volume fraction not occupied by the spheres, whereas the exponential factor describes the effect of the depletion layers. When ϕ is small, $\alpha \cong 1 - (1 + q)^3\phi$ and $4\pi(1 + q)^3R^3/3$ is the excluded volume connected with each sphere. As ϕ increases, the average distance between adjacent spheres decreases, and α becomes larger than $1 - (1 + q)^3\phi$ because the depletion layers overlap.

Because $c_2^{\rm II}/c_2^{\rm I} = \alpha(q, \phi^{\rm II})/\alpha(q, \phi^{\rm I})$, we substitute the expression for α provided by Eq. **5** into the ratio of the free volume fractions, and thereby we can determine an *apparent q* value for each experimental tie-line. We find that, within the experimental error, this quantity does not depend on the tie-line position in the coexistence curve. We present the average value of the apparent q as a function of the square root of PEG molecular weight as the filled circles in Fig. 3. We note that the q values calculated by using $\alpha = 1 - (1 + q)^3\phi$ are only ≈10% smaller than the values calculated by using the complete expression for α and so conclude that the overlapping of depletion layers contributes only marginally to the apparent q values reported in Fig. 3. We have also considered the effect of the PEG molecular weight polydispersity. Because the polydispersity of PEG is given by a narrow Poisson distribution function to be applied to PEG (38, 39), it can be shown that the polymer distribution in each of the two coexisting phases is essentially the same as in the original homogeneous system, and hence polydispersity does not significantly modify the slope of the tie-lines.

In Fig. 3, we also present the theoretical q values (open circles), calculated by using $q = 1.1R_g/R$. R_g for PEG is taken from ref. 15, whereas R is estimated from the molecular weight and specific volume of γD. As seen in Fig. 3, both theoretical q and the apparent q vary as the square root of the PEG molecular weight as expected for a Gaussian coil (40). However, the apparent values of q are numerically ≈50% smaller than the corresponding theoretical values. This significant discrepancy represents the deviation of the actual γD-PEG interactions relative to the sphere-ideal coil case.

It has been found that $q \approx R_g/R$ in the case of bovine serum albumin, chymotrypsinogen, and RNase, whereas $q \approx 0.5R_g/R$ in the case of β-lactoglobulin and lysozyme (12, 15). The difference between the values obtained for the various proteins suggests that the departure from the excluded volume model is related to direct interactions of PEG with protein surface residues (15). Our results clearly suggest the presence of such specific interactions in the γD-PEG-water systems. It appears that weak attractive interactions between PEG and γD induce a flattening of the polymer coils near the protein surface and a corresponding reduction of the depletion layer thickness.

The Spinodal Boundaries and the Dependence of the LLPS Temperature on PEG Concentration. The spinodal surface (40), which describes the spinodal temperature, $T_{\rm sp}$, as a function of protein and PEG concentrations, defines the boundary between the stable domain ($(\partial\hat{\mu}_1/\partial c_1)_{\hat{\mu}_2,T} > 0$) and the unstable domain ($(\partial\hat{\mu}_1/\partial c_1)_{\hat{\mu}_2,T} < 0$) of a homogeneous protein–PEG–water system. The spinodal condition, $(\partial\hat{\mu}_1/\partial c_1)_{\hat{\mu}_2,T{\rm sp}} = 0$, can be used to determine the spinodal temperature, $T_{\rm sp}$, as a function of c_1 and c_2. Indeed, on differentiating Eq. **3a** with respect to c_1 at constant $\hat{\mu}_2$ and T, we obtain:

$$\left(\frac{\partial\hat{\mu}_1}{\partial c_1}\right)_{\hat{\mu}_2,T} = \left(\frac{\partial\hat{\mu}_1'}{\partial c_1}\right)_T - \left[\left(\frac{\partial\xi}{\partial c_1}\right)_T^2 - \left(\frac{\partial^2\xi}{\partial c_1^2}\right)_T\right]c_2. \qquad [6]$$

In the case of hard sphere-ideal polymer coil excluded volume interactions, $\alpha = \exp(-\xi)$, and the difference $(\partial\xi/\partial c_1)_T^2 - (\partial^2\xi/\partial c_1^2)_T$ becomes equal to the quantity, $(\partial^2\alpha/\partial c_1^2)_T/\alpha$. If we use Eq. **5**, we find that, when both ϕ and q are small, $(\partial^2\alpha/\partial\phi^2)_T/\alpha$ is approximately equal to $12q^3$. Although Eq. **5** is inaccurate insofar as high-order derivatives such as $(\partial^2\alpha/\partial\phi^2)_T$ are concerned, it usefully describes the general features of the excluded volume model. Indeed, due to the overlapping of adjacent depletion layers, $(\partial^2\alpha/\partial\phi^2)_T$ must be positive and must increase as the depletion layer thickness increases. These general features of the excluded volume model can be used to understand the effect of PEG on initially stable protein aqueous solutions. In fact, because $(\partial\xi/\partial c_1)_T^2 - (\partial^2\xi/\partial c_1^2)_T$ is positive, Eq. **6** predicts that as the polymer concentration increases, $(\partial\hat{\mu}_1/\partial c_1)_{\hat{\mu}_2,T}$ decreases and can be made to reach zero, i.e., a spinodal boundary. Thus, sufficient PEG can in principle induce phase separation at any given temperature. This is true even if the protein–protein interactions are repulsive.

We now use the spinodal condition to analyze the dependence of the LLPS temperature on PEG concentration. We choose the spinodal condition because the spinodal surface is tangent to the coexistence surface on the critical line, and because it is math-

ematically more convenient than the coexistence condition. If we apply the spinodal condition to Eq. **6** and differentiate the second member with respect to both c_2 and T_{sp} at constant c_1, we obtain in the limit of c_2 approaching zero:

$$\left(\frac{\partial T_{sp}}{\partial c_2}\right)_{c_1} = \frac{(\partial\xi/\partial c_1)^2_{T_{sp}} - (\partial^2\xi/\partial c_1^2)_{T_{sp}}}{(\partial^2\hat{\mu}_1'/\partial c_1\partial T_{sp})}. \qquad [7]$$

The quantity $(\partial^2\hat{\mu}_1'/\partial c_1\partial T_{sp})$ is equal to $-(\partial^2 e'/\partial c_1^2)_{T_{sp}}/RT_{sp}^2$, where e' is the internal energy per unit volume of the protein–water binary system. On addition of another protein molecule to the system, the mean distance between the proteins must decrease, promoting more protein–protein interactions. From this second-order effect, it follows that the quantity $(\partial^2 e'/\partial c_1^2)_{T_{sp}}$ is negative for attractive protein–protein interactions, zero for hard-core interactions (i.e., hard spheres), and positive for repulsive interactions.

Eq. **7** correctly describes the sign of the dependence of the LLPS temperature as a function of PEG concentration for the γD-PEG-water systems. The LLPS temperature increases as the PEG concentration increases because of the attractive nature of the protein–protein interactions $[(\partial^2\hat{\mu}_1'/\partial c_1\partial T_{sp}) > 0]$. Also, $(\partial T_{sp}/\partial c_2)_{c_1}$ increases with PEG molecular weight because $(\partial\xi/\partial c_1)^2_T - (\partial^2\xi/\partial c_1^2)_T$ increases as the depletion layer thickness increases.

Conclusion

Our experimental findings and theoretical analysis demonstrate that liquid–liquid phase transition in ternary protein–PEG–water systems is a powerful tool for understanding the effect of PEG as a precipitating agent. We have shown that such studies can be used to locate the phase boundary for binary protein–water systems, which is very useful if this phase separation is hidden. We can also predict the change of protein solubility as a function of the concentration of crystallizing agents. Moreover, the slope of the tie-lines and the dependence of LLPS temperature on polymer concentration provide a sensitive check of the validity of excluded volume models for protein–polymer interactions. The increase of the LLPS temperature with PEG concentration is due to attractive protein–protein interactions.

We acknowledge Ajay Pande, Alexander Chernov, Annette Tardieu, Peter Vekilov, and Seth Fraden for useful critical comments. This work was supported by National Aeronautics and Space Administration Grant NAG8-1659 (to G.B.B.) and National Institutes of Health Grants EY05217 (to G.B.B.) and EY10535 (to J.P.).

1. Albertsson, P. A. (1986) *Partition of Cell Particles and Macromolecules* (Wiley, New York).
2. Abbott, N. L., Blankschtein, D. & Hatton, T. A. (1991) *Macromolecules* **24,** 4334–4348.
3. McPherson, A. (1999) *Crystallization of Biological Macromolecules* (Cold Spring Harbor Lab. Press, Plainview, NY).
4. Finet, S. & Tardieu, A. (2001) *J. Cryst. Growth* **232,** 40–49.
5. Kulkarni, A. M., Chatterjee, A. P., Schweizer, K. S. & Zukoski, C. F. (2000) *J. Chem. Phys.* **113,** 9863–9873.
6. Asakura, S. & Oosawa, F. (1954) *J. Chem. Phys.* **22,** 1255–1256.
7. Vrij, A. (1976) *Pure Appl. Chem.* **48,** 471–483.
8. Gast, A. P., Hall, C. K. & Russell, W. B. (1983) *J. Colloid Interface Sci.* **96,** 251–267.
9. Lekkerkerker, H. N. W., Poon, W. C. K., Pusey, P. N., Stroobants, A. & Warren, P. B. (1992) *Europhys. Lett.* **20,** 559–564.
10. Ilett, S. M., Orrock, A., Poon, W. C. K. & Pusey, P. N. (1995) *Phys. Rev. E* **51,** 1344–1352.
11. Atha, D. H. & Ingham, K. C. (1981) *J. Biol. Chem.* **256,** 12108–12117.
12. Arakawa, T. & Timasheff, S. N. (1985) *Biochemistry* **24,** 6756–6762.
13. Abbott, N. L., Blankschtein, D. & Hatton, T. A. (1992) *Macromolecules* **25,** 3932–3941.
14. Vergara, A., Paduano, L. & Sartorio, R. (2002) *Macromolecules* **35,** 1389–1398.
15. Bhat, R. & Timasheff, S. N. (1992) *Protein Sci.* **1,** 1133–1143.
16. Edmond, E. & Ogston, A. G. (1968) *Biochem. J.* **109,** 569–576.
17. Lomakin, A., Asherie, N. & Benedek, G. B. (1996) *J. Chem. Phys.* **104,** 1646–1656.
18. Malfois, M., Bonneté, F., Belloni, L. & Tardieu, A. (1996) *J. Chem. Phys.* **105,** 3290–3300.
19. Lomakin, A., Asherie, N. & Benedek, G. B. (1998) *Phys. Rev. Lett.* **77,** 4832–4835.
20. Asherie, N., Lomakin, A. & Benedek, G. B. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9465–9468.
21. Yu, M., Arons, J. S. & Smit, J. A. M. (1994) *J. Chem. Technol. Biotechnol.* **60,** 413–418.
22. ten Wolde, P. R. & Frenkel, D. (1997) *Science* **277,** 1975–1978.
23. Galkin, O. & Vekilov, P. G. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6277–6281.
24. Benedek, G. B. (1997) *Invest. Ophthalmolo. & Visual Sci.* **38,** 1911–1921.
25. Thomson, J. A., Schurtenberger, P., Thurston, G. M., Thomson, J. A. & Benedek, G. B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7079–7083.
26. Broide, M. L., Berland, C. R., Pande, J., Ogun, O. & Benedek, G. B. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5660–5664.
27. Chirgadze, Yu. N., Driessen, H. P. C., Wright, G., Slingsby, C., Hay, R. E. & Lindley, P. F. (1996) *Acta Crystallogr. D* **52,** 712–721.
28. McDermott, M. J., Gawinowicz-Kolks, M. A., Chiesa, R. & Spector, A. (1988) *Arch. Biochem. Biophys.* **262,** 609–619.
29. Asherie, N., Pande, J., Lomakin, A., Ogun, O., Hanson, S. R. A., Smith, J. B. & Benedek, G. B. (1998) *Biophys. Chem.* **75,** 213–227.
30. Schurtenberger, P., Chamberlin, R. A., Thurston, G. M., Thomson, J. A. & Benedek, G. B. (1989) *Phys. Rev. Lett.* **63,** 2064–2067.
31. Liu, C., Asherie, N., Lomakin, A., Pande, J., Ogun, O. & Benedek, G. B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 377–382.
32. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497.
33. Galkin, O., Chen, K., Nagel, R. L., Hirsch, R. E. & Vekilov, P. G. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 8479–8483.
34. Lekkerkerker, H. N. W. (1990) *Colloid Surf.* **51,** 419–426.
35. Tuinier, R., Vliegenthart, G. A. & Lekkerkerker, H. N. W. (2000) *J. Chem. Phys.* **105,** 10768–10775.
36. Widom, B. (1963) *J. Chem. Phys.* **39,** 2808–2812.
37. Lebowitz, J. L., Helfand, E. & Praestgaard, E. (1965) *J. Chem. Phys.* **43,** 774–779.
38. Flory, P. J. (1940) *J. Am. Chem. Soc.* **62,** 1561–1565.
39. Vergara, A., Paduano, L., Sartorio, R. & Vitagliano, V. (1999) *J. Phys. Chem. B* **103,** 8732–8738.
40. Tanford, C. (1961) *Physical Chemistry of Macromolecules* (Wiley, New York).