Mechanism of Precipitation of Proteins by Polyethylene Glycols

ANALYSIS IN TERMS OF EXCLUDED VOLUME*

Donald H. Atha‡ and Kenneth C. Ingham§

From the American Red Cross Plasma Derivatives Laboratory, Bethesda, Maryland 20014

(Received for publication, May 11, 1981)

The apparent solubilities of various proteins (14,000 to 670,000 daltons) were measured in the presence of polyethylene glycols (PEGs) of different sizes. All of the solubility curves, determined by measuring the protein concentration in the supernate of centrifuged mixtures, exhibited the characteristic linear dependence of log S (g/liter) on PEG concentration (w/v). For human albumin in PEG-4000 at pH 4.5, this linearity extended over a 1,000-fold range of solubility, even though the appearance of the sedimented phase changed from a viscous fluid to a white amorphous solid. The slope, \( \beta \), decreased from 0.27 to 0.09 with decreasing M, of PEG from 20,000 to 400, but was insensitive to changes in solution conditions (pH, T, salts), suggesting the absence of specific chemical interactions between protein and polymer. This conclusion was supported by the observation that concentrations of PEG up to 30% (w/v) had no significant effect on the melting temperature of ribonuclease A. Furthermore, equilibrium dialysis measurements, as well as various spectral measurements, provided no evidence for such interactions. Using a steric exclusion model (Edmond, E., and Ogston, A. G. (1968) Biochem. J. 109, 569-576) and assuming that the chemical potential of the solid phase is constant, \( \beta \) can be related to interaction coefficients calculated from co-volumes using the equivalent sphere radii of PEG (r2) and protein (r3). Although good agreement was obtained for albumin in PEG-20,000, the predicted dependence of \( \beta \) on r2 was greater than observed and the predicted dependence of \( \beta \) on r3 was of opposite direction to that observed. However, the interaction coefficient determined from the equilibrium dialysis measurements of albumin and PEG-1000 agreed with the predicted value. Thus, the exclusion of low concentrations of PEG by albumin can be explained by a simple excluded volume model, whereas the exclusion of protein out of solution by PEG appears to be more complex.

Polyethylene glycol is a nontoxic water-soluble synthetic polymer which is widely used in chemical and biomedical industries. In spite of its relatively inert chemical properties, PEG has been found to have some interesting biological effects. For example, it causes profound changes in the conformation of DNA polymers (Jordan et al., 1972), promotes the fusion and hybridization of cells (Davidson et al., 1976), and potentiates mitogen-induced stimulation of lymphocytes, presumably by enhancing the lateral aggregation of receptors in the membrane (Bessler et al., 1977). Our interest in PEG stems from its widespread use as a fractional precipitating agent for the purification of proteins from a variety of sources (Polson et al., 1964; Chun et al., 1967). We are attempting to develop its potential for the large-scale isolation of clinically useful proteins from human plasma. Because of its nondenaturing qualities, PEG provides an attractive alternative to ethanol for recovering minor components, many of which are difficult to obtain in native form from ethanol-generated fractions (Chandra and Wickerhauser, 1978; Wickerhauser et al., 1979; Hao et al., 1980).

The molecular basis of the protein-precipitating action of PEG and other polymers is poorly understood. The original work of Polson et al. (1964) documented the increasing effectiveness of PEG as the size of the polymer is increased. Jueckes (1971) subsequently called attention to the tendency of larger proteins to precipitate at lower concentrations of PEG and noted a correlation between the slope of the precipitation curve and the size of the protein. These trends which have since been confirmed in other laboratories (Honig and Kula, 1976; Kula et al., 1977; Ingham, 1977, 1978; Miekkä and Ingham, 1978) have reinforced the notion that the precipitation process is due primarily to excluded volume effects. According to this view, proteins are sterically excluded from regions of the solvent occupied by the inert synthetic polymers and are thus concentrated until their solubility is exceeded and precipitation occurs. Although this concept appears to provide a qualitative explanation for the existing observations, it has not been thoroughly explored on a quantitative level.

Studies with purified proteins indicate that the dependence of solubility, S, on PEG concentration, C, is strictly exponential with data generally conforming to the following equation (in addition to the present report, see Jueckes, 1971; Foster et al., 1973; Lee and Lee 1979; Middaugh et al., 1979)

\[
\log S = \beta C + \text{constant}
\]

The form of this equation is identical to that used to describe the solubility of proteins in the presence of inorganic salts (Cohn and Ferry, 1943). With salts, the linearity frequently holds only at high salt concentration; low concentrations may

* The abbreviations used are: PEG, polyethylene glycol, polyethylene oxide, and polyoxyethylene (chemical formula, \( \text{HOCH}_2\text{CH}_2\text{O(CH}_2\text{CH}_2\text{O)}_n\text{CH}_2\text{CH}_2\text{OH} \)); TEG, tetraethylene glycol; MPD, 2-methyl-2,4-pentanediol.

† Present Address, Massachusetts Institute of Technology, Whittaker College of Health Sciences, Technology, and Management, Cambridge, MA 02139.

‡ Supported by a career development award (HL 00325) from the National Heart, Lung and Blood Institute. To whom correspondence should be addressed.
actually increase solubility (Green, 1931; Melander and Horváth, 1977). By contrast, Middaugh et al. (1979) showed that with PEG, the linearity extends all the way to the ordinate, at least for those proteins whose intrinsic solubilities were sufficiently low to be accurately measured. They proposed that thermodynamic activities of more soluble proteins in saturated solution can be determined by extrapolation of solubility data obtained in the presence of PEG and subsequently utilized such an approach in an analysis of the thermodynamic basis for the abnormal solubility of cryoimmunoglobulins (Middaugh et al., 1980). The increasing interest in the behavior of proteins under conditions of high activity such as occur in some cells and organelles, or in the presence of polymer networks such as in the extracellular matrix of tissue, provides additional impetus for further investigation of protein-polymer mixtures. There is also considerable interest in the use of PEG for the production of protein crystals whose properties may be superior to those obtained by other means (Ward et al., 1975; McPherson, 1976; Eichele, et al., 1978; Grabowski et al., 1978).

In this report, we utilize expressions derived from the work of Ogston and co-workers (Ogston, 1958, 1962; Ogston and Phelps, 1961; Edmond and Ogston, 1966, 1970) to quantitatively evaluate the excluded volume concept as it applies to protein solubility in the presence of PEG. We have determined the solubility of various size proteins in a wide range of molecular weight polyethylene glycols in order to provide the necessary data for such an evaluation. In addition, we have measured the effect of PEG on the circular dichroic and absorption spectra and on the denaturation temperature of ribonuclease, and have conducted equilibrium dialysis experiments in the presence of albumin, in order to evaluate the extent to which PEG might interact with proteins, potentially complicating an analysis which relies completely on excluded volume considerations.

**EXPERIMENTAL PROCEDURES**

**Materials**—Human serum albumin (Grades A and B), bovine pancreatic α-chymotrypsin (Grade A), human γ-globulin (Fraction II, Grade B), human fibrinogen (Fraction I, Grade B), and bovine pancreatic ribonuclease (Grade A) were purchased from Calbiochem. Hen egg white lysozyme (Grade I), bovine α-lactalbumin (Grade II), rabbit muscle aldolase (Grade I), and bovine thyroglobulin (Type I) were from Sigma. Fractions of PEG having average molecular weights between 400 and 8000 were purchased from Fisher under the trade name Carbowax. PEG 20,000 was purchased from Hoechst. Tritium-labeled PEG-1,000 was obtained from New England Nuclear. Tetraethylene glycol was from Aldrich. 2-methyl-2,4-pentanediol was from Eastman.

**Spectroscopic Measurements**—The radius of the equivalent sphere, \( r_s \), was calculated using the relationship

\[
\frac{r_s^2}{V_s} = \frac{3 \times 10^{-6} \text{[g]} \cdot M_r}{4\pi N_A}
\]

where \( \sigma \) is Simha's factor (Simha, 1940; Mehl et al., 1940) = 0.5 for suspended spheres, \( N \) is Avogadro's number, and \( M_r \) is the average molecular weight (Table II). Intrinsically viscosities were calculated from the equation

\[
\eta_s \sim K \eta
\]

where \( K = 0.5 \) and \( \eta = 0.156 \) (ml/g) for the molecular weight range 200 to 8000 (Bailey and Koleske, 1967; Rossi and Cuniberti, 1964), and \( \eta = 0.78 \) and \( K = 0.0125\) (ml/g) for the molecular weight range 10^4 to 10^6 (Bailey et al., 1958). The length, \( l \), of a PEG fiber (0.3 mm/1000 dal) was calculated from CFP space filling models (Harte, 1969) based on known bond angles and lengths. The radius of the PEG fiber (\( R_F = 0.02 \) mm) was calculated from:

\[
R_F = M_r \cdot V_s / \pi N
\]

where \( V_s \) is the measured partial specific volume = 0.84 ml/g (average for PEGs 600 to 20,000) (Lepori and Mollica, 1978; Beje and Hvidt, 1972).

**Packing Density of Albumin**—An upper limit for the solubility of albumin can be estimated based on the maximum density of a hard sphere fluid (Hoover and Ree, 1968). The volume fraction occupied by a hard sphere is given by

\[
(\eta_p = 0.91)
\]

Absorbivity values (Å/mM) of 26.20, 20.4, 5.3, 9.1, 13.8, 15.5, and 10.4 were used for lysozyme, α-lactalbumin, chymotrypsin, human serum albumin, aldolase, γ-globulin, fibrinogen, and thyroglobulin, respectively. References are given in Table III. With concentrated solutions, the volume of the sedimented phase can become a significant fraction of the total volume. Since PEG is largely excluded from the solid phase, this can cause an increase in the concentration of PEG in the supernate (Middaugh et al. 1979). Most of the data reported here were obtained using bulk protein concentrations of 20 mg/ml or less, thus avoiding the need to correct for this effect.

**Equilibrium Dialysis Measurements**—Approximately 5 ml of buffered solutions containing human serum albumin at various concentrations were dialyzed using Spectropor 4 membrane tubing (M, cut-off, 12,000 to 14,000) against buffered solutions to which tritium-labeled PEG-1,000 had been added. The solutions were allowed to equilibrate at 23 °C in glass containers which were continuously inverted in a Kraft model RD250 rotator. Samples of 0.2-ml volume were vigorously mixed with 10 ml of scintillation fluid. After counting the entire series 6 times in a Packard model 3890 liquid scintillation spectrometer, the ratios of the counts inside to outside the bag were averaged. Control experiments involving the use of Protosol (New England Nuclear) or prior removal of protein by precipitation with trichloroacetic acid showed that the presence of protein inside the bag did not significantly affect the measured ratios. Comparison of samples in which the labeled PEG was added inside or outside the bag gave equivalent ratios, as did samples equilibrated 72 h versus 48 h. The final protein concentrations were determined optically using the absorbivity for albumin.

**Spectral Measurements**—Thermal denaturation of ribonuclease A was measured using a Perkin-Elmer MPF-4 spectrophotometer equipped with a thermostatted four-chamber cell holder and 0.5-cm diameter quartz cuvettes. The temperature was monitored using a Bailey model BAT9 thermosterm and thermocouple probe. The emission at 310 nm was monitored as a function of continuously increasing temperature, with excitation at 270 nm. Circular dichroism measurements were performed at 23 °C on a Cary model 61 spectropolarimeter with a thermostatted cell holder. Mean residue ellipticities, [\( \Theta \text{[res]} \), reported in deg cm²/dmol of residue, were calculated for ribonuclease A using a mean residue molecular weight of 115. The cell path length was 10 mm for the near and 0.2 mm for the far ultraviolet measurements. The protein concentration was determined using an absorbivity of 0.71 liter/cm g at 277 nm. Differential absorption measurements were made at 23 °C on a Cary 118C spectrophotometer using the tandem cell method (Herskovits, 1967; Donován, 1969).

**Molecular Parameters for Polyethylene Glycol**—The radius of the equivalent sphere, \( r_s \), was calculated using the relationship

\[
\frac{r_s^2}{V_s} = \frac{3 \times 10^{-6} \text{[g]} \cdot M_r}{4\pi N_A}
\]

where \( \sigma \) is Simha's factor (Simha, 1940; Mehl et al., 1940) = 0.5 for suspended spheres, \( N \) is Avogadro's number, and \( M_r \) is the average molecular weight (Table II). Intrinsically viscosities were calculated from the equation

\[
\eta_s \sim K \eta
\]

where \( K = 0.5 \) and \( \eta = 0.156 \) (ml/g) for the molecular weight range 200 to 8000 (Bailey and Koleske, 1967; Rossi and Cuniberti, 1964), and \( \eta = 0.78 \) and \( K = 0.0125\) (ml/g) for the molecular weight range 10^4 to 10^6 (Bailey et al., 1958). The length, \( l \), of a PEG fiber (0.3 mm/1000 dal) was calculated from CFP space filling models (Harte, 1969) based on known bond angles and lengths. The radius of the PEG fiber (\( R_F = 0.02 \) mm) was calculated from:

\[
R_F = M_r \cdot V_s / \pi N
\]

where \( V_s \) is the measured partial specific volume = 0.84 ml/g (average for PEGs 600 to 20,000) (Lepori and Mollica, 1978; Beje and Hvidt, 1972).

**Packing Density of Albumin**—An upper limit for the solubility of albumin can be estimated based on the maximum density of a hard sphere fluid (Hoover and Ree, 1968). The volume fraction occupied by a hard sphere is given by

\[
(\eta_p = 0.91)
\]
by hexagonally close packed hard spheres is 0.7405 (Scott, 1960). The maximum volume fraction for a hard sphere fluid, according to Hoover and Ree, is 0.667 times the volume fraction for hexagonal close packing, or 0.494. Using this value and a specific gravity of 1.37 g/cc for the albumin spheres, taken as the reciprocal of the partial specific volume (Peters, 1975; Oncley et al., 1947), an upper limit of 677 g/liter was obtained.

THEORETICAL SECTION

The chemical potentials of components in a ternary system can be expressed as functions of composition (Ogston, 1962; Edmond and Ogston, 1968). Simplified forms of the expressions for PEG (component 2) and protein (component 3), in which higher order virial coefficients are neglected, can be written as

\[ \mu_2 = \mu_2^0 + RT \ln (C_2 + cC_2 + aC_2) \quad 1a \]
\[ \mu_3 = \mu_3^0 + RT \ln (C_3 + dC_3 + aC_3) \quad 1b \]

In this treatment, \( C_2 \) and \( C_3 \) have units of moles/liter and the superscript zero refers to ideal solutions of pure components at unit molarity and standard conditions. The coefficients \( c, d, \) and \( a \) refer to PEG-PEG, protein-protein, and PEG-protein interactions, respectively, and have units of liters/mole. As the concentration of PEG is increased, the chemical potential of the protein increases, eventually exceeding that of the solid phase, at which point precipitation occurs. The following expressions are obtained from Equation 1b for the chemical potential of the protein under conditions of saturation in the absence \((\mu_3^0)\) and presence \((\mu_3)\) of PEG

\[ \mu_3 = \mu_3^0 + RT \ln (S_3 + dS_3) \quad 2a \]
\[ \mu_3 = \mu_3^0 + RT \ln (S_3 + dS_3 + aC_3) \quad 2b \]

where \( S_2 \) and \( S_3 \) represent the corresponding solubilities. Since \( \mu_2 \) (solution) = \( \mu_2 \) (solid), and assuming that the chemical potential of the amorphous solid phase is constant over the range of PEG concentration, we can equate 2a with 2b and obtain

\[ \ln S_2 = \ln S_3^0 + d (S_3 - S_0) - aC_2 \quad 3 \]

Under conditions where \( S_3 \) is small, the middle term can be neglected giving the required linear dependence of \( \ln S_2 \) on PEG concentration with a slope determined by the PEG-protein interaction coefficient and an intercept which reflects the intrinsic solubility in the absence of PEG. However, close inspection of Equation 3 suggests that even for highly soluble proteins such as albumin, there will be a concentration of polymer above which \( S_2 \ll S_3 \) and the middle term becomes constant

\[ \ln S_2 = \ln S_3^0 + dS_3 - aC_2 \quad 4 \]

The slope under such conditions will still be determined by \( a \), but the extrapolated intercept will reflect not only the solubility, \( S_3^0 \), but the protein-protein interaction coefficient, \( d \), as well. The product \( dS_3 \) relative to \( \ln S_3 \) is a measure of the nonideality of the saturated protein solution in the absence of PEG. In the units in which solubility data are presented in this study, \( \log (\text{grams/liter}) \) versus % PEG (w/v), \( a \) and \( d \) are given by the following expressions

\[ a = -0.23 \beta M_2 \]
\[ d = \frac{2.3 M_1}{S_3^0 \log (S_3^0)} \]

where \( \beta \) is the measured slope of the precipitation curve, \( M_2 \) is the average molecular weight of the PEG, \( M_1 \) is the molecular weight of the protein, \( S_3^0 \) (apparent) is obtained from the extrapolated intercept, and \( \gamma \) is the thermodynamic activity coefficient of the protein under saturating conditions in the absence of PEG.

RESULTS

Solvability Measurements

Albumin at pH 4.5—The solubility of albumin in the presence of PEG is very sensitive to pH with minimum solubility occurring at pH 4.5 near the isoelectric point (Ingham, 1978). Fig. 1 shows that the dependence of \( \log S \) on the concentration of PEG-4000 is linear over a 1000-fold range of solubility, even though the appearance of the sedimented phase changes from a viscous fluid to a white amorphous solid. The solid line is a least squares fit of the data from which a value of \( a = 180 \) liters/mol was calculated using Equation 5a. Linear extrapolation to the ordinate axis \((\text{dashed line})\) yields an intercept corresponding to an apparent solubility in the absence of PEG which, even under these isoelectric conditions, exceeds the absolute physical limit of solubility estimated from the inverse of the partial specific volume of the hydrated protein. Thus, variations in the second term in Equation 3 must become important at higher protein concentrations. Attempts to extend the measurements into this region were complicated by the extremely high solubility of albumin and the tendency of such concentrated solutions to form two aqueous phases upon addition of small amounts of PEG (Edmond and Ogston, 1968). An accurate value of \( S_3^0 \) would allow the protein-protein interaction coefficient, \( d \), to be calculated from the extrapolated intercept, using Equation 4. From our own experience, it can be said that \( S_3^0 \) is at least 400 g/liter under the conditions of Fig. 1 since clear solutions of this concentration can be readily prepared. As the concentration of albumin is further increased, the solutions become extremely viscous and gel-like, making it difficult to equilibrate and separate the solid and liquid phases. For the present purpose, we have chosen to estimate a realistic upper limit for the solubility of albumin based on a consideration of the packing density for hard spheres (see “Methods”). This approach yields \( S_3^0 = 677 \) g/liter for albumin which in turn leads to a lower limit of 151 liters/mol for the value of \( d \). If a value of 400 g/liter is used for the maximum solubility, the calculated value of \( d \) increases to 340 liters/mol. The dotted lines in Fig. 1 are theoretical curves generated with these two limiting values of \( d \).

Effect of Solution Conditions—The effects of varying the pH above or below 4.5 on the precipitation curve for albumin are shown in Fig. 2A. It is apparent that the horizontal position

---

1 The expressions given by Edmond and Ogston (1968) are in terms of molality. Our use of concentration in molarity facilitates comparison of observed interaction coefficients with those calculated from volumes in units of liters/mol. Data plotted on a w/v (molality) scale exhibit upward curvature to an extent which increases with increasing PEG concentration. However, in those cases where the solubility data extend to the ordinate (e.g. Fig. 2B), initial slopes are identical to those plotted on a w/v (molality) scale. The intriguing feature of the w/v plots is that they are linear over the entire range of concentrations, the slope at any point being equivalent to that which on a w/v scale is obtained only at the limit of low concentration.

2 Given that the appearance of the sedimented phase sometimes changes considerably over the range of PEG concentration, one might question the assumption of constant chemical potential. However, when supernates of different tubes were interchanged, vigorously mixed, and re-equilibrated, there was no significant change in the concentration of dissolved protein suggesting that such variation is not large. The consistent linearity of the solubility curves indicates that if \( \mu_{pept} \) varies, its incremental change/mol of PEG must be independent of PEG concentration, i.e., \( \frac{d\mu_{pept}}{d\text{PEG}} = \text{constant} \). The slope of the solubility curves would then be related to the difference between the interaction coefficients in the two phases.
The values of the slopes and extrapolated intercepts, together with the corresponding interaction coefficients calculated from them, are summarized in Table I. The range of values given for \( d \) was calculated on the assumption that 400 g/liter of the curve is much more sensitive to pH than is the slope. The values of the slopes and extrapolated intercepts, together with the corresponding interaction coefficients calculated from them, are summarized in Table I. The range of values given for \( d \) was calculated on the assumption that 400 g/liter \( \leq S_i \leq 677 \) g/liters as explained above. The much larger values at pH 7 versus pH 4.5 may be related to the large net charge (\(-18\) electrostatic unit at pH 7) which would be expected to increase repulsion between albumin molecules. The lower slope obtained at pH 7 may be related to the higher range of PEG concentrations required for measurements at this pH.

In the presence of 0.2 M ZnCl\(_2\), the solubility is drastically reduced and the pH dependence is much less pronounced. Because of the low value of \( S_i \), the middle term in Equation 3 never becomes significant and the curves are linear all the way to the ordinate. Thus, it is not possible to estimate \( d \) under these conditions. These effects are undoubtedly related to the binding of numerous zinc cations to the albumin molecule and the greater solubility at pH 4.5 relative to pH 7 is compatible with the fact that the isoelectric point of albumin is shifted to pH 7.0 in the presence of the metal ions (Schultz and Heremans, 1966). The slope is also insensitive to the presence of other salts such as KCl, KSCN, and potassium phosphate at 0.15 M, and to variations in temperature between \( 4^\circ \) and \( 30^\circ \) C.

Effect of PEG Size—The effectiveness of PEG in reducing

![Fig. 1. Effect of PEG-4000 on the solubility of human serum albumin. Measurements were made at 22-24 °C in 0.05 M potassium acetate buffer, pH 4.5, containing 0.1 M KCl. Total protein concentration. \( \Delta \), 125 mg/ml; \( \triangle \), 20 mg/ml; \( \bigotimes \), 5 mg/ml; \( \odot \), 2 mg/ml. The solid line is a least squares fit of the data to a straight line. The dashed line gives a linear extrapolation of solubility in the absence of PEG. The upper dotted line was calculated according to Equation 3 where \( a = 180 \) liters/mol, \( d = 151 \) liters/mol, and \( S_i = 677 \) g/liters. The lower dotted line was calculated using \( a = 180 \) liters/mol, \( d = 340 \) liters/mol and \( S_i = 400 \) g/liter. \( 1/\bar{Y} \) is the maximum theoretical solubility based on the partial specific volume of the hydrated protein (Onckley et al., 1947).

![Fig. 2. Effect of pH on the solubility of human serum albumin and thyroglobulin in PEG. A, effect of pH on the solubility of human serum albumin in PEG-4000. Data indicated by filled symbols were obtained with a total protein concentration of 20 mg/ml in the following buffers: 0.05 M potassium acetate, 0.1 M KCl, pH 3.8-5.2, and 0.05 M potassium phosphate, 0.1 M KCl, pH 7.0. Measurements in the presence of ZnCl\(_2\) (open symbols) were made at a total protein concentration of 5 mg/ml in the following buffers: 0.05 M potassium acetate, 0.1 M KCl, 0.2 M ZnCl\(_2\), pH 4.5, and 0.05 M Tris, 0.1 M KCl, 0.1 M ZnCl\(_2\), pH 7.0. The dashed line is taken from Fig. 1. B, effect of molecular weight of PEG on the solubility of bovine thyroglobulin at neutral and acid pH. Measurements were made at a total protein concentration of 20 mg/ml in the following buffers: 0.05 M potassium acetate, 0.1 M KCl, pH 4.5 and 4.6; 0.08 M potassium phosphate, 0.1 M KCl, pH 7.0.

![Fig. 3. Effect of molecular weight of PEG on the solubility of human serum albumin. Measurements were made in 0.05 M potassium acetate buffer, pH 4.5, containing 0.1 M KCl. Total protein concentration was 20 mg/ml, except for PEG-4000 data, which is described in Fig. 1. The solid lines for PEG-600, -600, -4000, and -4000 are least squares fits. The lines for PEG-6000, PEG-20,000, and TEG were arbitrarily drawn through the common intercept.]

<table>
<thead>
<tr>
<th>pH</th>
<th>Slope, ( \beta )</th>
<th>Intercept</th>
<th>Interaction Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.8</td>
<td>(-0.16)</td>
<td>3.92</td>
<td>120</td>
</tr>
<tr>
<td>4.5</td>
<td>(-0.23)</td>
<td>3.50</td>
<td>180</td>
</tr>
<tr>
<td>5.2</td>
<td>(-0.21)</td>
<td>3.67</td>
<td>161</td>
</tr>
<tr>
<td>7.0</td>
<td>(-0.15)</td>
<td>5.10</td>
<td>113</td>
</tr>
<tr>
<td>7.0 (+ZnCl(_2))</td>
<td>(-0.21)</td>
<td>0.54</td>
<td>161</td>
</tr>
<tr>
<td>4.5 (+ZnCl(_2))</td>
<td>(-0.21)</td>
<td>1.0</td>
<td>161</td>
</tr>
</tbody>
</table>
the solubility of proteins increases with increasing size of the polymer. Data for human albumin at pH 4.5 are given in Fig. 3. There is a progressive shift in the position of the solubility curve to lower concentration of PEG as the molecular weight is increased from 194 (tetraethylene glycol) to 20,000. There is also a steady increase in the slope which appears to approach a maximum value at a $M_w \approx 6,000$.

The values of the slopes and intercepts obtained by least squares analysis of the data, along with the corresponding interaction parameters, are summarized in Table II. The value of $-0.28$ for PEG-6000 is in excellent agreement with that obtained by Juckes (1971) for bovine albumin under similar conditions. Within experimental error, the intercepts are constant between PEG-400 and PEG-4000, with small decreases for the larger polymers. However, as shown in Fig. 3, a consistent description of the data can be obtained by assuming that all of the curves intersect at a common point and that changes in polymer size affect only the slope. This is reasonable since neither $S_s$ nor $d$ should depend on the type of PEG.

Additional data for a much larger protein, thyroglobulin, are shown in Fig. 2B. At pH 4.5-4.6, the solubility curves are linear all the way to the ordinate axis. Under these conditions, where $S_s$ is sufficiently small that the middle term in Equation 3 is negligible, there is still a substantial increase in the slope upon increasing the molecular weight from 400 to 4000. At neutral pH where the solubility is much greater, the slope for PEG-4000 is essentially the same as at pH 4.6. However, the extrapolated intercept under these conditions (not shown) corresponds to an apparent solubility in excess of $1/5$ indicating the importance of protein-protein interactions.

**Effect of Protein Size**—The solubilities of several proteins in PEG-4000 under constant solution conditions are shown in Fig. 4. All of the proteins exhibit the characteristic linear dependence of log solubility on PEG concentration. Lysozyme and $\alpha$-lactalbumin, which are highly homologous and have similar structures, showed similar behavior under these conditions, even though their isoelectric points differ greatly (4.7 for albumin versus 10.5 for lysozyme). Although the larger proteins (fibrinogen, $\gamma$-globulin, aldolase, and thyroglobulin) tend to precipitate at lower PEG concentrations than the smaller ones (lysozyme, $\alpha$-lactalbumin, and albumin), there is not a good correlation between the position of the solubility curve and the size of the protein. However, the slopes of the curves exhibit a definite increase with increasing protein size. This is illustrated in Fig. 5A where the observed slopes are plotted against the equivalent-sphere radii, calculated for each protein from published diffusion constants using the Stokes

| TABLE II |

Effect of PEG size on slopes and interaction coefficients for human serum albumin

Slopes and intercepts were obtained by a least squares fit of the data in Fig. 3. Observed protein-PEG interaction coefficients were determined from the slopes using Equation 5a. Calculated protein-PEG interaction coefficients were obtained for the sphere and fiber models of PEG using Equations 5a, 7c, and 8a, and the molecular dimensions given in Tables II and III.

<table>
<thead>
<tr>
<th>PEG</th>
<th>$M_a$</th>
<th>$r_s$</th>
<th>Slope, $\beta$</th>
<th>Intercept</th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEG</td>
<td>194</td>
<td>0.64</td>
<td>0.98</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.60</td>
<td>-0.02</td>
<td>3.44</td>
<td>8.2</td>
<td>176</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>0.72</td>
<td>-0.11</td>
<td>3.58</td>
<td>15</td>
<td>192</td>
<td>280</td>
<td></td>
</tr>
<tr>
<td>1,000</td>
<td>0.94</td>
<td>-0.14</td>
<td>3.41</td>
<td>32</td>
<td>224</td>
<td>378</td>
<td></td>
</tr>
<tr>
<td>2,000</td>
<td>1.16</td>
<td>-0.23</td>
<td>3.50</td>
<td>80</td>
<td>363</td>
<td>552</td>
<td></td>
</tr>
<tr>
<td>4,000</td>
<td>1.72</td>
<td>-0.28</td>
<td>3.16</td>
<td>512</td>
<td>593</td>
<td>2690</td>
<td></td>
</tr>
<tr>
<td>6,000</td>
<td>2.65</td>
<td>-0.27</td>
<td>2.84</td>
<td>1244</td>
<td>1331</td>
<td>5055</td>
<td></td>
</tr>
<tr>
<td>8,000</td>
<td>3.46</td>
<td>-0.26</td>
<td>2.58</td>
<td>2092</td>
<td>2092</td>
<td>5055</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Number average molecular weights, $M_a$, for PEG-400-6000 were taken as the midpoints of the range of values appearing in the Union Carbide Corporation Trade Literature for Carbowax Poly(ethylene glycols) (1978). The value for Hoescht PEG-20,000 was assumed.

$^b$ The radius of an equivalent sphere, $r_s$, was calculated from intrinsic viscosities (see "Methods").
equation (see Table III). The values of the slopes and the corresponding values of \(a\) are also given in Table III.

### Protein-Polymer Interactions

In this section, we present the results of experiments designed to evaluate the extent to which PEG might interact with proteins. Such interactions, if they occur, would obviously be important in understanding the mechanism of precipitation.

**Thermal Denaturation of Ribonuclease**—Ribonuclease undergoes a well characterized temperature-dependent unfolding in acid or in the presence of urea and other denaturants (Hermans and Scheraga, 1961; Brandts and Hunt, 1966). A number of studies have appeared in which the effects of various cosolvents on the melting temperature of ribonuclease and other proteins were used to learn about the interactions of those substances with the protein and about their influence on water structure (see review by Franks and Eagland, 1975).

If PEG interacts appreciably with ribonuclease and if the extent of that interaction is different for the native and denatured forms, then the presence of the polymer might be expected to shift the melting temperature. Similarly, if PEG altered the water structure near the surface of the protein, one might also expect a shift.

The effects of temperature on the fluorescence intensity of ribonuclease in the presence of various concentrations of PEG-400 and PEG-4000 at pH 2.6 are shown in Fig. 6. In all cases, there is an abrupt transition characterized by a substantial increase in fluorescence intensity as buried nonfluorescent tyrosines are exposed to the aqueous environment (Cowgill, 1966). It is apparent that neither PEG-4000 nor PEG-400 had a significant effect on the melting temperature, even at concentrations as high as 30% (w/v). Under these same conditions, ethanol at a concentration of 25% lowered the midpoint of the transition by -8 °C (data not shown; see Brandts and Hunt, 1966).

**Circular Dichroism Measurements**—The circular dichroic spectra of ribonuclease in the absence and presence of 10% PEG-4000 are shown in Fig. 7, A and B. The solution conditions are identical to those used in the thermal denaturation experiments. No significant effects of PEG could be seen in either the near or far UV regions, indicating that the secondary and tertiary structures of ribonuclease are not significantly perturbed (Greenfield and Fasman, 1969; Strickland, 1974). This is consistent with measurements of the circular dichroic spectra of tubulin by Lee and Lee (1979) which showed only small differences in the presence of 10% PEG.

**Differential Absorption Measurements**—Addition of organic solvents to aqueous solutions of ribonuclease perturbs the absorption spectrum of the surface tyrosyl residues (Herskovitz, 1967; Donovan, 1969). The magnitude of the perturbation relative to that observed with free tyrosine is frequently used to learn about the extent to which tyrosyl residues are exposed to the solvent and therefore capable of responding to the altered dielectric properties. Alternatively, it might be possible to use a well characterized protein like ribonuclease as a probe to examine the solvent composition in the immediate domain of the protein molecule. In a previous study (Ingham, 1977) it was shown that the effectiveness of PEG as a spectral perturbant increased with increasing molecular weight. Most of the effect occurred in the range up to \(M_w = 400\) with a smaller increase between \(M_w = 400\) and 4000. Fig. 7C presents the difference spectrum obtained for ribonuclease at neutral pH in 20% PEG-400. Also shown for comparison is a spectrum obtained in the presence of 2-methylpentanediol at the same concentration. The latter compound was the subject of a detailed study by Pitz and Timasheff (1978), who reported that ribonuclease is preferentially hydrated in solutions containing MPD. They concluded that MPD molecules are effectively excluded from the surface of the protein and have a low probability of forming direct contacts with the nonpolar residues even at MPD concentrations as high as 50% (v/v). The observations in Fig. 7C suggest that both PEG and MPD are quite capable of altering the dielectric environment of the surface tyrosines, a result which seems incompatible with the notion of total exclusion from the surface of the protein.
Exclusion of Proteins by Polyethylene Glycols

Equilibrium Dialysis Measurements—Mixtures of human albumin and $[^{15}N]$PEG-1000 were dialyzed extensively with appropriate controls to ensure that equilibrium was attained (see "Methods"). Under these conditions, the chemical potential of PEG (component 2) is the same on both sides of the membrane and the following expression can be derived from Equation 1b

$$\ln \frac{C_2}{C_2^0} = c \left( \frac{C_2}{C_1} - C_2 \right) - a C_2$$

where $C_2$ and $C_2^0$ are the concentrations of PEG inside and outside the bag, respectively. Since extremely low concentrations of PEG were employed, the middle term involving $c$ should be negligible and a plot of $\log C_2/C_2^0$ versus protein concentration ($C_2$) should yield a straight line with slope determined by $a$. The results shown in Fig. 8 indicate that the concentration of PEG inside the dialysis bag decreases exponentially with increasing protein concentration. At the highest protein concentration employed (274 mg/ml), the ratio of $C_2/C_1$ was 0.42. The slope of the best fit line in Fig. 8 corresponds to a value of $a = 210$ liters/mol which is $\sim 7$-fold greater than the corresponding value of 32 obtained from the slope of the albumin solubility curve in Fig. 3. However, it is in excellent agreement with the theoretical value of 224 calculated on the basis of excluded volume considerations using an equivalent-sphere model for PEG (see below).

Excluded Volume Considerations

So far, our treatment has been independent of any consideration of the molecular mechanism by which PEG lowers the solubility of proteins. The coefficients $a$ and $d$ whose values are reported above, are empirically derived quantities whose thermodynamic significance relates to their appearance in Equations 1 to 4 and 6. However, their mechanistic significance is a separate question. The results of the previous section reinforce the concept of PEG as a benign reagent with little tendency to interact with proteins and thus lend support to the concept that its precipitating action stems primarily from excluded volume effects. The most extensive attempt to quantitatively apply excluded volume concepts to the interpretation of protein solubility in the presence of PEG was made by Juckes (1971), who also called attention to the validity of the salting out equation in this context and reported that the slope, $B$, increased approximately linearly with increasing Stokes radius of the proteins. However, the number and size range of globular proteins included in that study were somewhat limited and the inclusion of a large virus particle in the series might not be appropriate. Furthermore, the effect of the molecular weight of PEG was not addressed. In this section we examine the quantitative validity of the simplest form of the excluded volume model by comparing observed slopes and interaction coefficients with values calculated on the basis of that model.

In the absence of specific interactions, the coefficients $a$, $c$, and $d$ can be regarded as expressing the volumes mutually excluded by neighboring molecules (Flory, 1953). In their analysis of the sedimentation, osmotic pressure, and phase-separation behavior of ternary aqueous systems, Edmond and Ogston (1968) considered the simplified view that the molecules could be treated as equivalent impenetrable spheres with co-volumes given by

$$U_{12} = 10^5 c = \frac{4\pi N}{3} \left(2r_2 \right)^3$$

$$U_{13} = 10^5 d = \frac{4\pi N}{3} \left(2r_3 \right)^3$$

$$U_{33} = 10^5 r = \frac{4\pi N}{3} \left(r_3 + r_3 \right)^3$$

where $U_{12}$, $U_{13}$, and $U_{33}$ are the molar excluded covolumes for pairs of spherical molecules 2-2, 2-3, and 3-3 with equivalent radii $r_2$ and $r_3$ (expressed in cm) and $N$ is Avogadro's number. If one views the PEG as a fiber with radius $R_F$ and length $l$, while continuing to regard the protein molecules as spheres, then (Edmond and Ogston, 1968; Ogston, 1970)

$$U_{23} = 10^5 a = N\pi \left(R_F + r_3 \right)^3 \left(2r_2 \right)^3$$
The molecular parameters which we have used for the calculation of interaction coefficients are summarized in Tables II and III. Calculated values of \( \alpha \) for albumin as a function of PEG molecular weight are also given in Table II where they can be readily compared with the "observed" values determined from the solubility measurements. Both the sphere-sphere and sphere-fiber models correctly predict the increase in \( \alpha \) with increasing PEG size. However, quantitative agreement between theory and experiment is limited to the higher molecular weight polymers and then only for the sphere-sphere model. Furthermore, as shown in Fig. 5B, the observed decrease in slope with decreasing polymer size is not reproduced; both models predict a trend with decreasing molecular weight of PEG which is not only steeper but in the opposite direction to that which is observed. Similar results were obtained with thyroglobulin (not shown). These discrepancies are too large to be accounted for by uncertainties in the molecular weights of the PEGs.

The predicted influence of protein size on the slope of the precipitation curve is also at variance with the experimental observations. The sphere-sphere model predicts a slope which is proportional to the cube of the covolume radius given by the sum \( r_2 + r_1 \) in Equation 7c. As illustrated by the solid line in Fig. 5A, the discrepancy between the observed and theoretical values for PEG-4000 increases with increasing radius of the protein. Slopes calculated on the basis of the sphere-fiber model were too large to appear on the figure. Finally, the tendency of the observed values to be concave towards the abscissa is opposite to that predicted by either model.

A similar approach can also be used to estimate the contribution of excluded volume effects to the protein-protein interaction coefficient, \( d \). Using a value of 3.52 nm for the radius of albumin, Equation 7b yields \( d = 880 \) liters/mol. This is within the range of "observed" values determined for albumin at neutral pH but is substantially larger than those determined at lower pH (Table I). The discrepancy is opposite to that which would be expected from considerations of net charge and suggests the occurrence of attractive interactions near the isoelectric point which would counteract the effects of excluded volume.

**DISCUSSION**

The most striking feature of the precipitation of proteins with PEG is the consistent linearity of the semilog plots of protein solubility versus PEG concentration. This linearity would seem to justify the neglect of higher virial coefficients in the thermodynamic expressions describing this phenomenon and indicates that the PEG-protein interaction parameter \( \alpha \) as defined in Equation 1b is invariant over an extremely broad range of concentrations of both macromolecules. Furthermore, \( \alpha \) is relatively insensitive to changes in solution conditions including pH, temperature, and ionic strength up to 0.15 M. This suggests that chemical interactions involving attractive or repulsive forces between PEG and protein are relatively unimportant in the precipitation mechanism. The absence of a significant effect of temperature on the slope of the solubility curve is consistent with a mechanism of precipitation that is primarily entropic, as is an excluded volume mechanism. The failure of high concentrations of PEG to have a measurable effect on the CD spectrum or melting temperature of ribonuclease provides additional support for the inert nature of this polymer and suggests that it does not interact with the protein. This is in agreement with the measurements of Amiconi et al. (1977) which showed that the sensitive functional properties of ligand binding to hemoglobin are largely unaffected by polyethylene glycol. Further evidence for the lack of an appreciable attraction between PEG and protein comes from published observations that the rate of escape of PEG-4000 through a microporous ultrafiltration membrane was unaffected by the presence of albumin between 1 and 100 mg/ml (Busby and Ingraham, 1980). The same publication also showed that the elution of \[^{14}C\]PEG-4000 from a Sephadex G-75 column was not significantly affected by the excess albumin in the applied sample.

The linearity of log \( S \) versus PEG plots has been exploited by Middaugh et al. (1979, 1980) to estimate, by extrapolation, the thermodynamic activity of saturated protein solutions. This is an intriguing approach which is rendered all the more attractive by the experimental difficulties of making direct measurements on concentrated protein solutions. Under conditions of low solubility, the extrapolated intercept provides a measure of the actual solubility, \( S' \). Under conditions of high solubility, the intercept, \( S'' \) (apparent), exceeds \( S' \) (actual) by an amount which depends on the protein-protein interaction coefficient, \( d \), and which is related to the activity coefficient, \( \gamma \), under saturating conditions (Equation 5b). Using the values of \( d \) given for albumin in Table I, we estimate values of \( \gamma \) ranging from 4.7-7.9 at pH 4.5 to 186-314 at pH 7. The values for albumin at pH 4.5 are in reasonable agreement with the value of 4 estimated by Middaugh et al. (1979) for hemoglobin at pH 7 (near its isoelectric point). These authors commented on the surprising degree of ideality and called attention to the 100-fold higher values of \( \gamma \) determined by Ross and Minton (1977) from analysis of osmotic pressure and sedimentation equilibrium data obtained under similar conditions. A decision regarding the validity of thermodynamic activities determined from solubility measurements would best be made following a more extensive comparison with values obtained by other methods under identical conditions.

The observed effects of protein and polymer size on the slopes of the precipitation curves are not adequately explained by the simple excluded volume model. Although reasonable agreement was obtained between the observed and calculated slopes for albumin in the presence of the higher molecular weight PEGs, this agreement did not extend to larger proteins or to smaller PEGs. Nevertheless, there is a clear dependence on macromolecular size which, when combined with the lack of specific chemical interactions, seems qualitatively compatible with the notion of a steric exclusion mechanism. Although the accuracy of the geometrical parameters used to calculate covolumes can be quite arbitrary and arbitrary variations of those parameters within reasonable limits does not improve the situation. For example, according to the simple covolume model, as the size of the PEG becomes negligible compared to that of the protein, the interaction coefficient (Equation 7c) approaches a constant but the calculated slope continues to vary inversely with the molecular weight (Equation 5a). Even when PEG-400 is allowed to have zero radius, the calculated slope, which then constitutes a lower limit for the excluded volume theory, is substantially larger than the observed one.

It should be emphasized that the covolume treatment is strictly valid only in dilute solution (Flory, 1953). At higher concentrations, the amount of volume excluded per molecule is expected to diminish because of the overlap between the covolume radii (Tanford, 1961). The neglect of this effect gives rise to predicted interaction coefficients which are unrealistically large. However, it is not clear how such effects can be incorporated into the analysis without introducing higher virial terms into Equation 1b, which would destroy the requisite linearity. In those cases where the solubility is sufficiently low to be measured in the absence of PEG (see Fig. 2, as well as the work of Middaugh et al., 1979), the slope is constant all the way to the ordinate. Thus, if we chose to do
so, we could confine our analysis to dilute solutions of PEG and protein in order to circumvent any objections to utilizing a dilute solution theory to analyze data, most of which are obtained at high concentration. It is important to emphasize that whatever model is used to estimate the value of \( \alpha \) must take into account the fact that it is a unique parameter, characteristic of a particular protein-polymer pair, and independent of protein and polymer concentration.

Using the PEG radii in Table II, one can calculate the fraction of the total solution volume which would be excluded by PEG molecules as a function of concentration, assuming no overlapping or shrinking of the equivalent spheres. Alternatively, one can consider the volume actually occupied by the molecules themselves as estimated from the partial specific volume. According to the latter picture, addition of PEG of any molecular weight to a concentration of 50% would cause less than a two-fold decrease in the available volume. At the other extreme, an 8% solution of 20,000-dalton equivalent spheres would have no volume left for protein molecules unless they were sufficiently small to penetrate the interior spaces of the random coil polymer. As the concentration of PEG increases, the polymers may either shrink and become more compact or interpenetrate forming a gel-like network. In the latter picture, the \( \alpha \) coefficient would be related to the dependence of the size of the interior spaces of the network on polymer concentration.

The equilibrium dialysis measurements with [\(^{3}H\)PEG-1000] provide an alternate approach to the measurement of \( \alpha \) which is not complicated by potential artifacts arising from the use of high PEG concentrations or by assumptions of constant chemical potential of protein in the solid phase. It is therefore of interest that the value of \( \alpha \) determined in this fashion, although much greater than that determined from solubility measurements, is in excellent agreement with that predicted from the osmotic coefficient in the presence of PEG and observed the characteristic linear dependence of \( \log S \) on PEG concentration with slopes of 0.16, 0.20, and 0.28 for PEG-1000, -4000, and -6000. Using these slopes, we calculate respective values of \( \alpha = 0.57, 0.54 \), and 0.56 liters/mol. These are much smaller than the corresponding values of 384, 578, and 884 liters/mol calculated by Equation 7c using a value of 4.4 nm for the equivalent sphere radius of the tubulin dimer (Frigon and Timasheff, 1975). These calculated values are lower limits since polymeric species could have been present under the conditions employed (Frigon and Timasheff, 1975). Lee and Lee also made densitometric measurements on mixtures which had been extensively dialyzed and determined values for the change in protein chemical potential with increasing PEG concentration, a parameter which is closely related to \( \alpha \), since, by Equation 1b, \( \alpha = (\partial \mu / \partial \phi C)/RT \). Using values from their Fig. 4A (at PEG concentrations corresponding to the midpoints of their solubility curves), we calculate \( \alpha = 260 \) liters/mol for PEG-1000 and 470 liters/mol for PEG-4000. These values are much larger than those based on solubility and are much closer to the values predicted from osmotic work. Thus, "preferential hydration" of tubulin in solutions of PEG can be largely explained on the basis of a simple excluded volume effect, whereas precipitation of tubulin cannot.

The implication of the present study with respect to the use of PEG in fractional precipitation deserves comment. Since all of the solubility data presented here were obtained with purified proteins, it will be of interest to determine the extent to which the observed parameters can be duplicated in mixtures. A recent report from this laboratory pointed out that protein-protein interactions can have rather profound effects on solubility in the presence of PEG but that nonspecific interactions such as those arising from electrostatic attraction between oppositely charged proteins can be minimized by maintaining ionic strength at physiological or higher levels (Miekka and Ingham, 1980). Furthermore, the total protein concentration in most mixtures of interest is sufficiently low that protein-protein interactions due to excluded volume effects can be neglected. In the absence of specific interactions, the order of precipitation of proteins from a mixture upon addition of PEG would then be expected to depend on two factors: the relative values of \( \alpha \) as determined primarily by size, and the initial concentrations of the respective proteins relative to their apparent solubilities under the given solution conditions. A large protein with a steep slope which is present at relatively low concentration may precipitate later than a small one which precipitates with a shallow slope but whose initial concentration is closer to its solubility limit. Manipulation of solution conditions can be expected to improve the separation of a given pair of proteins only to the extent that their intrinsic solubilities diverge. PEG can be viewed as an inert solvent sponge which indiscriminately raises the effective concentration (i.e. activity) of all of the proteins, those of larger size being somewhat more sensitive than smaller ones.

Acknowledgments—We are grateful to Dr. A. G. Ogston, Dr. Allen P. Minto, and Dr. Petre Munk for many helpful suggestions. We thank Dr. Shirley I. Miekka for useful discussions and for providing us with solubility data on \( \alpha \)-chymotrypsin. We also thank Dr. Jim Osborne for the use of the CD spectropolarimeter, Ronald Ling, Marcella Wozniak, Shelles Brev, and Kenneth Edwards for technical assistance, and Mareha Coté for secretarial assistance.

REFERENCES
Bailey, F. E., Kucera, J. L., and Imhof, L. G. (1958) J. Polymer Sci. 32, 517-538
Boje, L. and Hvidt, A. (1972) Biopolymers 2, 2357-2364
Foster, P. R., Dunnill, P., and Lilly, M. D. (1973) Biochim. Biophys. Acta 317, 505-516
Frigon, R. P., and Timasheff, S. N. (1975) Biochemistry 14, 4559-
Exclusion of Proteins by Polyethylene Glycols


Green, A. A. (1931) J. Biol. Chem. 93, 495–542


Rossi, C., and Cuniberti, C. (1964) Polymer Letters 2, 681


Schwert, G. W. (1951) J. Biol. Chem. 190, 799–806


