Lawrence Berkeley National Laboratory
Recent Work

Title
van der Waals Theory of Precipitation for Solutions of Binary Globular-Protein Mixtures

Permalink
https://escholarship.org/uc/item/7n2873vw

Author
Hino, T.

Publication Date
1998-07-01
van der Waals Theory of Precipitation for Solutions of Binary Globular-Protein Mixtures

Toshiaki Hino and John M. Prausnitz

Chemical Sciences Division

July 1998
Submitted to AIChE Journal
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
van der Waals Theory of Precipitation for Solutions of Binary Globular-Protein Mixtures

Toshiaki Hino and John M. Prausnitz
Department of Chemical Engineering
University of California
and
Chemical Sciences Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720, U.S.A.
July 1998

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.
van der Waals Theory of Precipitation for Solutions of Binary Globular-Protein Mixtures

Toshiaki Hino and John M. Prausnitz*

Department of Chemical Engineering
University of California, Berkeley
and
Chemical Sciences Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

Abstract

A theory of the van der Waals form is used to establish solid-fluid phase diagrams for aqueous solutions containing two kinds of globular proteins. Theory is based on a one-fluid model that uses a composition-dependent hard-sphere diameter and a composition-dependent solvent-mediated effective temperature to represent protein solutions. The hard-sphere reference equation of state is based on the model of Young that correlates computer simulations of fluid-solid coexistence curves for binary hard-sphere mixtures. The attractive perturbation term uses the inverse-power potential with variable exponent $n$. For single-protein systems, and for large $n$, our model predicts the observed metastable fluid-fluid coexistence curve that lies below the stable fluid-solid freezing line; the theoretical phase diagram with $n=7$ agrees with aqueous solubility data for lysozyme. Based on the correlation of George and Wilson and Rosenbaum et al., the optimum range of effective temperature is determined for protein crystallization. For binary protein systems, as the size difference increases, the mutual solubility of unlike proteins declines sharply in the precipitated solid phase. Illustrative plots are presented to indicate how the two-protein phase diagram depends on mixture composition, solubility difference, effective temperature, and interaction between unlike proteins. The theoretical results obtained here may be useful for design of protein-separation processes by crystallization.

(Key words: equation of state; protein crystallization; protein mixtures.)

* to whom correspondence should be addressed
Introduction

In polymer solutions, the polymeric solute can be separated by inducing phase separation. The precipitated phase is either a solid or a dense fluid phase rich in polymer. Similarly, a protein can be separated from an aqueous solution by precipitation, often achieved by using inorganic salts or nonionic polymers as precipitating agents. In the biotechnology industry, protein precipitation is widely used to recover and purify proteins from aqueous solutions (Rothstein, 1994). In addition, protein crystallization has been extensively studied by biologists performing X-ray crystallography of protein crystals (Rosenberger, 1996). However, phase diagrams of protein solutions remain less explored than those of solutions containing common polymers.

Protein solutions are similar to colloidal solutions. To understand the phase behavior of protein solutions, it is necessary to consider both the fluid-fluid phase separation that leads to two fluid phases and the fluid-solid phase separation that results in a fluid phase and a solid phase (Asherie et al., 1996; Löwen, 1997). Recent theoretical and experimental studies show that, in a temperature-density diagram of colloidal solutions, the fluid-fluid coexistence curve lies below the fluid-solid freezing line when the range of attractive interaction is sufficiently small relative to the particle diameter. (Asherie et al., 1996; Berland et al., 1992; Bolhuis and Frenkel, 1994; Broide et al., 1996; Daanoun et al., 1994; Hagen and Frenkel, 1994; Ilett et al., 1995; Lekkerkerker et al., 1992; Löwen, 1997; Muschol and Rosenberger, 1997; Poon, 1997; Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996; Tanaka et al., 1997; Tavares and Sandler, 1997).

In systems where the fluid-fluid coexistence curve lies below the fluid-solid freezing line, it is the fluid-solid phase separation that leads to the global minimum of the Gibbs energy. Fluid-fluid phase separation is metastable with respect to fluid-solid phase separation because the former only leads to a local minimum of the Gibbs energy (Poon, 1997). Under careful experimental conditions, however, fluid-fluid phase separation can be observed by light scattering before crystallization proceeds (Broide et al., 1991; Broide et al., 1996; Muschol and
Rosenberger, 1997; Taratuta et al., 1990; Thomson et al., 1987). Similarly, metastable fluid-fluid phase separation is also possible for solutions of crystallizable polymers (Burghardt, 1989).

In protein crystallization, the existence of fluid-fluid phase separation has an important consequence on the morphology of the protein precipitate (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996; ten Wolde and Frenkel, 1997). When proteins are precipitated at effective temperatures where fluid-fluid phase separation is possible, fluid-fluid phase separation interferes with crystallization that proceeds much more slowly than fluid-fluid phase separation. Because fluid-fluid phase separation favors large density fluctuations, amorphous solids are produced that are not suitable for X-ray diffraction analysis (Poon, 1997). To precipitate proteins as crystals, moderately poor solvents are desirable, poor enough to precipitate proteins but not too poor to induce fluid-fluid phase separation (George and Wilson, 1994; Poon, 1997; Rosenbaum et al., 1996).

With few exceptions, previous theoretical work has been confined to single-protein solutions. In this work, we are concerned with development of a molecular-thermodynamic model to compute phase diagrams for ternary solutions, that is, those that contain two types of globular proteins in addition to the solvent. We refer to these as binary solutions. The present work is stimulated by experimental results reported by Judge et al. (1995) who demonstrated recovery and purification of ovalbumin from an aqueous mixture of ovalbumin, conalbumin, and lysozyme by salt-induced bulk crystallization. Their study indicates that bulk crystallization may be feasible for the recovery of a target protein from aqueous multicomponent protein mixtures. In addition, we are also motivated by recent advances in modeling the phase behavior of single-protein systems and those for computing fluid-solid equilibria of binary hard-sphere mixtures.

This work presents an elementary van der Waals-type equation-of-state theory. For our reference equation of state, we use the model of Young (1993) that correlates the computer simulation of fluid-solid coexistence curves for binary hard-sphere mixtures (Kranendonk and Frenkel, 1991). For our perturbation term, we use an inverse-power (Sutherland) potential with variable exponent $n$ combined with the model of Daanoun et al. (1994) for a solid with
coordination number \( z \). In a temperature-density diagram for sufficiently large \( n \), our model predicts the fluid-fluid coexistence curve that lies below the fluid-solid freezing line.

In our model, the solid phase is, in general, a protein mixture. In previous theoretical work for polymer-induced precipitation of protein mixtures, Mahadevan and Hall (1990, 1992) assumed that only one type of protein precipitates in the solid phase except at the eutectic point where two pure solid phases and a fluid phase coexist. For a fluid phase that contains two types of proteins, Mahadevan and Hall (1990, 1992) used a theoretical two-component model.

We use a much simpler one-component model. As is common in mixture thermodynamics, we use the one-fluid approximation: we obtain the equation of state for protein mixtures by semi-empirically extending the equation of state for single-protein systems using composition-averaged equation-of-state parameters. Here, "one-component" refers to the solute. The solvent is a continuous medium. We use the mean-field perturbation theory that is simpler than the first-order perturbation theory used by Mahadevan and Hall (1990, 1992).

We begin with single-protein systems and determine a suitable exponent for the inverse-power potential by correlating aqueous lysozyme solubility data. Next, we establish the optimum range of effective temperature for protein crystallization. Finally, for solutions of protein mixtures, we compute phase diagrams as functions of mixture composition, size difference, solubility difference, effective temperature, and interaction between unlike proteins.

**Theoretical Framework**

For modeling protein solutions with a one-component model, we use the hard-sphere diameter and effective temperature that represent, respectively, the size of globular protein and solvent-mediated interaction between proteins. In a one-component model, where the solvent is a medium, the equation of state gives the osmotic pressure of the solution. For solutions containing two kinds of proteins, equation-of-state parameters are averaged over (solvent-free) mixture composition using appropriate mixing rules.
In protein precipitation by inorganic salts or nonionic polymers, the effective temperature depends on several experimental variables such as absolute temperature, ionic strength, pH, salt type, and polymer molecular weight. For most typical condensed systems, external pressure has little effect unless the pressure is very large. Accurate one-component models are not yet available for quantitatively computing the effective temperature as a function of all experimental variables. Therefore, our discussion is given in terms of the effective temperature that, at least in principle, can be obtained from osmotic-second-virial-coefficient measurements. Such measurements can be made as a function of temperature, pH, salt type, and ionic strength (Coen et al., 1995; Curtis et al., 1998; Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996).

**Pair Potential**

We assume that globular proteins in solution interact through the pair potential $\mathcal{u}(\tilde{R})$ for hard-sphere systems with an attractive inverse-power potential of variable exponent $n$:

$$
\mathcal{u}^{\text{IP}}(\tilde{R}) = \begin{cases} 
\infty & \tilde{R} < 1 \\
-\varepsilon\tilde{R}^{-n} & (n > 3) \\ 
-\varepsilon & 1 \leq \tilde{R} 
\end{cases}
$$

(1)

where $\tilde{R}=R/\sigma$ is the reduced center-to-center distance of spheres ($R=$center-to-center distance of spheres; $\sigma=$hard-sphere diameter); $\varepsilon$ is the potential energy parameter that specifies the strength of solvent-mediated attractive interaction between proteins; and superscript IP denotes inverse-power potential. In eq. (1), parameter $n$ systematically varies the range of attractive interaction relative to the hard-sphere diameter. As $n$ increases, the reduced range of attractive interaction declines. We use the inverse-power potential because of its simplicity.

For the inverse-power potential, the second virial coefficient is given by

$$
B_2^{\text{IP}} = 2\pi\sigma^3 \int_0^{\infty} \left[ 1 - e^{-\mathcal{u}^{\text{IP}}(\tilde{R})/k_BT} \right] \tilde{R}^2 d\tilde{R} = \frac{2\pi\sigma^3}{3} \left[ 1 - \frac{3}{n-3} \frac{1}{T} \right]
$$

(2)
where $k_B$ is the Boltzmann constant; $T$ is the absolute temperature; and $\bar{T}$ is the effective reduced temperature for the inverse-power potential given by

$$\bar{T} = \frac{k_B T}{\epsilon}.$$  \hspace{1cm} (3)

In the colloid and protein literature, $k_B T/\epsilon$ is often called the effective temperature. In our notation, $\epsilon/k_B$ is the effective temperature and $k_B T/\epsilon$ is the effective reduced temperature.

For salt solutions, parameter $\epsilon$ depends on pH, salt type, and ionic strength. For solutions containing a nonionic polymer, $\epsilon$ depends on polymer type, polymer molecular weight, and polymer concentration. Parameter $\epsilon$ may also depend on the absolute temperature because aqueous protein solutions involve specific interactions such as hydrogen bonding whose effects are strongly dependent on the absolute temperature. Smaller effective temperatures represent more attractive interactions between proteins.

**Helmholtz Energy and Equation of State**

In this work, the Helmholtz energy is given by a hard-sphere reference equation of state and a perturbation term that represent, respectively, the repulsive interaction for the hard-sphere system and attractive interaction for the inverse-power potential:

$$\left(\frac{A}{Nk_B^2T}\right) = \left(\frac{A}{Nk_B^2T}\right)_{\text{ref}} + \left(\frac{A}{Nk_B^2T}\right)_{\text{pert}}$$  \hspace{1cm} (4)

where $A$ is the Helmholtz energy and $N$ is the number of molecules. The equation of state is related to the Helmholtz energy by differentiating with respect to the number density:

$$\left(\frac{pV}{Nk_B^2T}\right) = p \left[\frac{\partial(A/Nk_B^2T)}{\partial \rho}\right]_{T, N} = \left(\frac{pV}{Nk_B^2T}\right)_{\text{ref}} + \left(\frac{pV}{Nk_B^2T}\right)_{\text{pert}}$$  \hspace{1cm} (5)
where $p$ is the osmotic pressure; $V$ is the volume; and $\rho = N/V$ is the number density. In eqs. (4) and (5), subscripts ref and pert denote reference and perturbation, respectively.

**Hard-Sphere Reference Equations**

For the reference Helmholtz energy, we use the model of Young (1993). This model is convenient because it gives analytic equations for correlations based on computer simulations of fluid-solid coexistence curves both in the one-component hard-sphere system and in binary hard-sphere mixtures (Carnahan and Starling, 1969; Hoover and Ree, 1968; Kranendonk and Frenkel, 1991; Young, 1993).

Consider first binary mixtures of hard spheres wherein the mole fraction of component 1 is $x$ and the hard-sphere diameter of component $i$ is $\sigma_i$. By choosing component 1 as the hard sphere with larger diameter, we define parameter $\alpha$ as the ratio of smaller-to-larger hard-sphere diameters:

$$\alpha = \frac{\sigma_{22}}{\sigma_{11}} \quad (\alpha \leq 1) . \quad (6)$$

In a one-component model, the packing fraction (i.e., reduced density) of mixture $y$ is given in terms of the average hard-sphere diameter:

$$y = \frac{\pi \rho \sigma^3}{6} . \quad (7)$$

Here, $\sigma$ is a composition-averaged hard-sphere diameter given by

$$\sigma^3 = x^2 \sigma^3_{11} + 2x(1-x)\sigma^3_{12} + (1-x)^2 \sigma^3_{22} \quad (8)$$
\[ \sigma_{12} = \frac{1}{2}(\sigma_{11} + \sigma_{22}) \]  \hspace{1cm} (9)

When \( \sigma^3 \) is given by eq. (8), for mixtures with \( \alpha \neq 1 \), eq. (7) is not equal to the "real" packing fraction of the mixture \( \pi \rho [x_1 \sigma_{11}^3 + x_2 \sigma_{22}^3] / 6 \). For mixtures, eq. (7) gives a pseudo packing fraction.

**Fluid Phase.** For the fluid phase, the reference hard-sphere Helmholtz energy and equation of state are given by (Young, 1993)

\[
\left( \frac{A}{N k_B T} \right)_{\text{ref}} = \frac{4y - 3y^2}{(1 - y)^2} + \ln \frac{6y}{\pi} + x \ln x + (1 - x) \ln (1 - x) \]  \hspace{1cm} (10)

\[
\left( \frac{pV}{N k_B T} \right)_{\text{ref}} = \frac{1 + y + y^2 - y^3}{(1 - y)^3} \]  \hspace{1cm} (11)

For one-component systems, eq. (11) is the Carnahan-Starling equation of state for the hard-sphere fluid (Carnahan and Starling, 1969).

**Solid Phase.** For the solid phase, the hard-sphere reference Helmholtz energy and equation of state are given by (Young, 1993)

\[
\left( \frac{A}{N k_B T} \right)_{\text{ref}} = -3 \ln \frac{V^* - 1}{V^*} + 5.124 \ln V^* - 20.78 V^* + 9.52 V^* - 5.95 V^* / 3 \]
\[ + 15.022 + x \ln x + (1 - x) \ln (1 - x) \]  \hspace{1cm} (12)

\[
\left( \frac{pV}{N k_B T} \right)_{\text{ref}} = \frac{3}{V^* - 1} + 2.566 + 0.55(V^* - 1) - 1.19(V^* - 1)^2 + 5.95(V^* - 1)^3 \]  \hspace{1cm} (13)
where

\[ V^* = \frac{V}{V_0}; \quad V_0 = N\sigma_0^3 \sqrt[3]{2} \]  \hspace{1cm} (14)

\[ \sigma_0^3 = x^2 \sigma_{11}^3 + 2f(\alpha)x(1-x)\sigma_{12}^3 + (1-x)^2 \sigma_{22}^3 \]  \hspace{1cm} (15)

where

\[ f(\alpha) = 1 + 13.5(1-\alpha)^{2.5} \]  \hspace{1cm} (16)

Equations (12) and (13) are based on the fit to the computer-generated compressibility factor for a one-component hard-sphere solid (Hoover and Ree, 1968; Young, 1993). Equation (12) contains a constant 15.022. This constant is necessary to give the correct melting point of one-component hard-sphere systems determined by computer simulations (Hoover and Ree, 1968; Young, 1993).

For mixtures, the essential part of the model of Young is function \( f(\alpha) \) given by eq. (16). Young determined \( f(\alpha) \) such that the model approximately fits the computer-generated fluid-solid coexistence curves for binary hard-sphere mixtures in the range \( 0.85 \leq \alpha \leq 1 \) (Kranendonk and Frenkel, 1991; Young, 1993). In the solid phase, the computer simulation of Kranendonk and Frenkel (1991) shows that the mutual solubilities of hard-sphere mixtures decline sharply as \( \alpha \) decreases below about 0.9.

**Perturbation Terms**

The perturbation term represents the solvent-mediated attractive interaction between proteins whose magnitude is represented (inversely) by reduced effective temperature \( \tilde{T} \). Using a one-fluid mixing rule, the reduced effective temperature of the mixture (at constant \( T \)) is given by
\[ \tilde{T}^{-1} = x^2 \tilde{T}_{11}^{-1} + 2x(1-x)\tilde{T}_{12}^{-1} + (1-x)^2 \tilde{T}_{22}^{-1} \]  

(17)

where

\[ \tilde{T}_{ij} = \frac{k_B T}{\varepsilon_{ij}} \]  

(18)

\[ \varepsilon_{12} = (1-k_{12})\sqrt{\varepsilon_{11}\varepsilon_{22}} \]  

(19)

where \( k_{12} \) is an adjustable parameter that relates the magnitude of attractive interaction between unlike proteins to those between like proteins.

Similar to parameter \( \alpha \) defined by eq. (6), we also define parameter \( \beta \) as the ratio of \( \varepsilon_{22} \) to \( \varepsilon_{11} \):

\[ \beta = \varepsilon_{22}/\varepsilon_{11} \]  

(20 a)

At constant absolute temperature

\[ \beta = \tilde{T}_{11}/\tilde{T}_{22} \]  

(20 b)

**Fluid Phase.** For the fluid-phase perturbation term, we use the van der Waals perturbation combined with the inverse-power potential given by eq. (1). For the Helmholtz energy, the van der Waals perturbation term is proportional to the product of number density and the average of the pair potential evaluated over the center-to-center intermolecular distance beyond the hard-sphere diameter. The equation of state is obtained from the Helmholtz energy.

For the fluid phase, the perturbation term for the van der Waals Helmholtz energy and that for the equation of state are identical, given by
where $y$ is the pseudo packing fraction of the fluid phase given by eq. (7).

**Solid Phase.** For the solid-phase perturbation term, we follow the model of Daanoun et al. (1994) that represents the solid phase by a perfect crystal wherein the intermolecular separation is determined by the ratio of the system density to the close-packed density. In this model, the perturbation term for the Helmholtz energy is proportional to the product of lattice coordination number $z$ and the pair potential evaluated at the system's intermolecular separation. Daanoun et al. considered only single-component crystals. We introduce a one-solid, analogous to a one-fluid, mixing rule for mixtures.

For the solid phase, the perturbation term for the Helmholtz energy and that for the equation of state are given by (Daanoun et al., 1994)

\[
\left( \frac{a}{Nk_BT} \right)_{pert} = - \frac{12}{n-3} \frac{y}{\bar{T}} 
\]

(21)

\[
\left( \frac{pV}{Nk_BT} \right)_{pert} = - \frac{12}{n-3} \frac{y}{\bar{T}} 
\]

(22)

where $u$ is the pair potential and $\bar{R}$ is the average reduced distance between closest neighbors in the lattice. $\bar{R}$ is given by

\[
\bar{R} = \left( \frac{y_{cp}}{y} \right)^{1/3} 
\]

(25)
where $y_{cp}$ is the close-packed packing fraction and $y$ is the pseudo packing fraction of the solid phase given by eq. (7).

In computing the perturbation term for the solid phase, it is a crude assumption to impose a particular structure on the solid phase. It is more appropriate to consider parameters $z$ and $y_{cp}$ as adjustable parameters whose values are limited by physical significance. Thus, $z$ should lie somewhere between 6 and 12 and $y_{cp}$ between 0.52 and 0.74. For our purposes here, we assume that the solid phase is the face-centered-cubic (fcc) lattice with $z=12$ and $y_{cp}=0.74$. We leave parameter $n$ as an adjustable parameter. The other adjustable parameter is $k_{12}$. Parameter $\sigma_n$ is estimated from crystallographic data and parameter $\varepsilon_n$ is determined from solubility or osmotic-second-virial-coefficient data.

**Computation Procedure**

In computing the fluid-solid coexistence curves in one-component models, we assume that the effective temperature of each component in the fluid phase is equal to that in the solid phase. In other words, we assume a uniform solvent composition throughout the (protein-free) fluid and solid phases. Calculations based on this simplifying assumption for one-component models are able to correlate observed solubility curves for several colloidal solutions (Poon, 1997; Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996).

**Single-Protein Systems**

For single-protein systems, we compute reduced effective temperature-density diagrams by equating the osmotic pressure and chemical potential of protein $\mu$ of coexisting phases at the same temperature. The chemical potential of protein is defined by

$$\mu = \left. \frac{\partial A}{\partial N} \right|_{T, \nu}$$

Equation (26) applies to both the fluid phase and the solid phase.
**Binary-Protein Systems**

For aqueous solutions containing two kinds of proteins, the phase diagram is calculated by equating the osmotic pressure and, for each protein, chemical potential $\mu_i$ of coexisting phases at the same temperature. The chemical potential of component $i$ is defined by

$$\mu_i \equiv \left( \frac{\partial A}{\partial N_i} \right)_{T, V, N_{\text{tot}}}$$

(27)

where $N_i$ is the number of molecules of component $i$. Equation (27) applies to both the fluid phase and the solid phase.

For binary-protein systems, we compute osmotic pressure-(solvent-free) composition diagrams at a fixed absolute temperature. In computing a fluid-solid coexistence curve and a solid-solid coexistence curve, we first fix the osmotic pressure or one of two solvent-free compositions of coexisting phases. Next, using the initial guesses for unknown variables, the reduced densities of coexisting phases are calculated from the equations of state. Iterations are repeated until the equality of chemical potential of coexisting phases is satisfied for each component.

As shown later for some binary-protein systems, there is a eutectic point where a fluid phase is in equilibrium with two solid phases having different compositions. For a eutectic system, the osmotic pressure and three solvent-free compositions of coexisting phases are simultaneously determined by equating, for each component, chemical potential of coexisting phases.

**Results and Discussion**

As pointed out by numerous authors, in modeling the interaction between colloidal particles in solutions, the important parameter is the one that specifies the range of attractive interaction relative to the particle diameter (Asherie et al., 1996; Bolhuis and Frenkel, 1994; Daanoun et al., 1994; Hagen and Frenkel, 1994; Ilett et al., 1995; Löwen, 1997; Poon, 1997;
In the inverse-power potential given by eq. (1), parameter \( n \) determines the range of attractive interaction relative to the particle diameter. The reduced range of attractive interaction declines with increasing \( n \).

We first show that as \( n \) rises, our model predicts a temperature-density diagram wherein the fluid-fluid coexistence curve lies below the fluid-solid freezing line. Next, we determine parameter \( n \) by correlating aqueous solubility data for lysozyme. We also calculate the optimum range of reduced effective temperature for protein crystallization, called the crystallization window or slot. Finally, we compute phase diagrams for aqueous protein mixtures as functions of mixture composition \( (x) \), size difference \( (\alpha) \), potential-energy difference \( (\beta) \), effective temperature \( (\tilde{T}) \), and interaction between unlike proteins \( (k_{12}) \). Parameter \( \beta \) indicates solubility difference. For a fixed absolute temperature and for globular proteins of similar size, if \( \epsilon_2 > \epsilon_1 \), the solubility of protein 2 is less than that of protein 1.

In the following figures, the fluid and solid regions of the phase diagram are denoted by \( F \) and \( S \), respectively. Similarly, the fluid-solid two-phase regions are denoted by \( F-S \) or by \( S-F \).

**Single-Protein Systems**

*Effect of the Reduced Range of Attractive Interaction.* For single-protein systems, parts a and b of Figure 1 show the theoretical temperature-density diagrams for \( n=5 \) and 6, respectively. In the perturbation term, we assume that the solid phase is the face-centered-cubic (fcc) lattice with \( z=12 \) and \( y_{cp}=0.740 \). These figures show that parameter \( n \) has a significant effect on the location of the fluid-fluid coexistence curve relative to the fluid-solid freezing line. As the reduced range of attractive interaction declines with increasing \( n \), theory predicts a phase diagram wherein the fluid-fluid coexistence curve is buried inside the fluid-solid domain.

At temperatures where both fluid-fluid and fluid-solid phase separations are possible (part b of Figure 1), only the fluid-solid phase separation leads to a global minimum of the Gibbs energy of the mixture. For the system in part b of Figure 1, fluid-fluid phase separation gives
two metastable fluid phases that lead to a local minimum of the Gibbs energy of the mixture (Poon, 1997).

In the present model, there is no stable fluid-fluid phase separation for \( n \) greater than about 5.5. However, both \( z \) and \( y_{cp} \) in eqs. (23) to (25) affect the critical value for \( n \) that separates Figure 1a-type phase behavior from Figure-1b type phase behavior. In this work, we do not adjust \( z \) and \( y_{cp} \) but use only parameter \( n \) as a system-independent adjustable parameter determined from the observed phase diagram for a single-protein system.

**Comparison with Experiment.** To compare theory with experiment, we need the hard-sphere diameter and solvent-mediated effective temperature. For a reasonable approximation, the hard-sphere diameter is estimated from X-ray analysis of protein crystals. The effective temperature is obtained from osmotic-second-virial-coefficient measurements. However, the theoretical osmotic second virial coefficient depends on both hard-sphere diameter and reduced effective temperature. Therefore, to obtain better agreement of theory with experiment, it is advantageous to regress simultaneously the hard-sphere diameter and effective temperature using additional solubility data. To fit experimental solubility data for a given system, the effective temperature depends on the choice of hard-sphere diameter.

Despite their simplicity, one-component models combined with two-parameter pair potentials have been successfully used to model phase diagrams of colloidal solutions, including those that contain globular proteins. Recently, Rosenbaum et al. and Poon showed that, in terms of packing fraction (i.e., reduced density) and reduced effective temperature, observed phase diagrams of several colloidal solutions follow closely the theoretical phase diagram for the adhesive hard-sphere potential (Poon, 1997; Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996). These results suggest that, in terms of reduced variables, there is a universal phase diagram for colloidal solutions. These results imply that a corresponding-states description may be applicable to protein solutions.
We compare our model with aqueous solubility data for lysozyme reported by Rosenbaum et al. (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996) who presented solubility data in terms of $\tau$, the reduced temperature for the adhesive hard-sphere potential. In the appendix, we discuss the procedure to express their data in terms of the effective temperature for our model given by eq. (3).

Figure 2 shows the phase diagram for aqueous lysozyme at pH 4.6. The open and solid circles show solubility data of Rosenbaum et al. (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996) calculated with $\sigma=3.4$ nm and $\sigma=3.2$ nm, respectively. (Rosenbaum et al. obtained effective temperature $\tau$ from osmotic-second-virial-coefficient data. Data shown by open and solid circles are calculated from the same set of experimental data. Figure 2 also shows the reduced phase-separation temperatures associated with the metastable fluid-fluid phase separation measured by Taratuta et al. (1990); they are given by open and solid triangles for $\sigma=3.4$ nm and $\sigma=3.2$ nm, respectively (Rosenbaum and Zukoski, 1996). In Figure 2, the solid and broken curves are the theoretical fluid-solid and fluid-fluid coexistence curves, respectively, for the inverse-power potential with $n=7$ and $z=12$.

Solubility data calculated with $\sigma=3.2$ nm agree with the fluid branch of the theoretical coexistence curve for our model with $n=7$. On the other hand, solubility data calculated with $\sigma=3.4$ nm follow the fluid branch of the theoretical coexistence curve for the adhesive hard-sphere potential presented by Rosenbaum and Zukoski (1996). In Figure 2, the primary difference between these models is in the high-density region. This discrepancy may be due to the mean-field perturbation theory used in our model. The fluid-solid lines shown by Rosenbaum et al. (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996) are those obtained by computer simulation (Hagen and Frenkel, 1994) for the attractive Yukawa potential. In the high-density region of our model, the theoretical densities of coexisting fluid and solid phases can be increased by adjusting $z$ and $\gamma_{cp}$.

For the metastable fluid-fluid phase separation, the critical solution temperature for our model with $n=7$ is close to that for the adhesive hard-sphere potential obtained from the Percus-
Yevick approximation for the hard-sphere radial distribution function (Watts et al., 1971). Therefore, for lysozyme, both models overestimate the critical solution temperature associated with the fluid-fluid phase separation. However, Rosenbaum and Zukoski (1996) also report that for γ\textsubscript{II}-crystallin, the observed solubility curve and critical solution temperature agree with the prediction by the adhesive hard-sphere potential.

Figure 2 shows semi-quantitative agreement of theory with experiment. In the present model, the solubility data for globular proteins are represented by the theoretical phase diagram for n=7 and z=12.

**Crystallization Slot.** The significant result from the work of Rosenbaum et al. (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996) is that the phase diagrams of different colloidal solutions may be represented by a universal phase diagram when the phase diagrams are constructed in terms of reduced density and reduced effective temperature. Consequently, the effective temperature determined from osmotic second virial coefficient measurements may be used to identify the optimum condition for protein crystallization called the crystallization window or slot. The work by Rosenbaum et al. provides support for the earlier work by George and Wilson (1994) who proposed to identify favorable conditions for protein crystallization from osmotic-second-virial-coefficient measurements.

For the adhesive hard-sphere potential, the crystallization slot is approximately given by 0.06<τ<0.15 (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996). Assuming that the hard-sphere diameter for the adhesive hard-sphere potential is equal to that for our model, the crystallization slot can also be expressed in terms of our reduced effective temperature \( \tilde{T} \) by converting τ to \( \tilde{T} \) using eq. (A 3).

For the present model, the crystallization slot is approximately given by

\[ 0.18 < \tilde{T} < 0.45 \] 

(for \( n=7 \) and \( z=12 \)).
Next, we compute the phase diagrams of protein mixtures at reduced temperatures in the crystallization slot.

**Aqueous Mixtures of Proteins**

For aqueous mixtures of globular proteins, we present reduced osmotic pressure-composition diagrams for several sets of parameters $\alpha$, $\beta$, and $k_{12}$ defined, respectively, by eqs. (6), (20), and (19). Phase diagrams at high osmotic pressures correspond to those in the high-reduced-density region. In the following figures, $x$ is the (solvent-free) mole fraction of component 1 in the binary mixture of proteins. All calculations are made at effective reduced temperatures where fluid-fluid phase separation does not exist.

$\alpha=1$ and $\beta \neq 1$. Part a of Figure 3 shows theoretical phase diagrams for $\alpha=1$, $\beta>1$, and $k_{12}=0$ at $\tilde{T}_{11}=0.45$. In this figure, there is no difference in the hard-sphere diameters of components 1 and 2. For $\beta>1$, component 1 has larger solubility than component 2. Therefore, the precipitated solid phase is rich in component 2. As $\beta$ increases, the composition difference between fluid and solid phases rises at a given composition of the fluid phase.

Part b of Figure 3 shows the effect of $k_{12}$ on the theoretical diagram for $\alpha=1$ and $\beta=1.2$ at $\tilde{T}_{11}=0.45$. Between unlike proteins, positive $k_{12}$ represents more unfavorable interaction than that with $k_{12}=0$. Conversely, negative $k_{12}$ represents more favorable interaction between unlike proteins than that with $k_{12}=0$. At a given composition of the fluid phase, the composition difference between fluid and solid phases for $k_{12}=0.08$ is larger than that for $k_{12}=0$ and $\beta=1.2$ in part a of Figure 3. For $k_{12}=-0.2$, the composition difference between fluid and solid phases is smaller than that for $k_{12}=0$ and $\beta=1.2$ in part a of Figure 3.

$\alpha \neq 1$ and $\beta = 1$. To elucidate the effect of size difference between unlike proteins, parts a, b, and c of Figure 4 show theoretical phase diagrams for $\alpha=0.95$, 0.9, and 0.85, respectively, with $\beta=1$ and $k_{12}=0$ at $\tilde{T}_{11}=0.45$. For the system in part a with $\alpha=0.95$, there is little composition...
difference between the liquid and solid phases. In the solid phase, the mutual solubility between unlike proteins decreases sharply as the size parameter $\alpha$ decreases below about 0.9.

For the system in part c, there is a eutectic point where two solid phases and a fluid phase coexist. At osmotic pressures below that at the eutectic point, essentially only one type of protein precipitates in the solid phase. The type of solid precipitate depends on the overall composition of the mixture. At osmotic pressures above that at the eutectic point (i.e., at high protein densities), two almost immiscible solid phases coexist.

For aqueous binary mixtures of globular proteins, the disparity in size between unlike proteins has a significant effect on the partitioning of protein mixtures between coexisting fluid and solid phases. In fluid-fluid phase separations for $\alpha > 0.85$ at sufficiently small $\tilde{T}_{11}$, however, there is no significant size effect on the partitioning of protein mixtures between coexisting fluid phases. For mixtures of globular proteins with sufficient disparity in size, sharp separation of a target protein may be best achieved through fluid-solid precipitation, not by inducing fluid-fluid phase separation.

When there is favorable interaction between unlike proteins, we expect that the effect of size disparity declines. In the present model, favorable interaction between unlike proteins is represented by negative $k_{12}$. Parts a and b of Figure 5 show theoretical phase diagrams for $k_{12} = -0.1$ and $k_{12} = -0.2$, respectively, with $\alpha = 0.85$ and $\beta = 1$ at $\tilde{T}_{11} = 0.45$. As expected for decreasing $k_{12}$, in part b of Figure 5, the magnitude of the size effect that favors phase separation declines and the composition difference between coexisting phases decreases relative to that in part a of Figure 5.

Variation of Effective Reduced Temperature. We next examine how effective reduced temperature affects the phase diagrams of protein mixtures at fixed parameters $\alpha$, $\beta$, and $k_{12}$. As the effective reduced temperature decreases, the solubility of globular proteins declines. We now discuss how the partitioning of protein mixtures between coexisting fluid and solid phases varies with the effective reduced temperature.
Parts a, b, and c of Figure 6 show theoretical phase diagrams for $\tilde{T}_{11}=0.45$, $\tilde{T}_{11}=0.4$, and $\tilde{T}_{11}=0.35$, respectively, with $\alpha=0.9$, $\beta=1$, and $k_{12}=0$. As the effective reduced temperature declines, the composition difference between coexisting phases increases and theory predicts solid-solid phase separation in the high-density region of the phase diagram.

Finally, we examine the system where both $\alpha$ and $\beta$ differ from unity. Parts a and b of Figure 7 show theoretical phase diagrams for $\beta=1$ and $\beta=0.9$, respectively, with $\alpha=0.9$ and $k_{12}=0$ at $\tilde{T}_{11}=0.35$.

Because the solubility of pure component 2 increases as $\beta$ declines, at a given composition of the fluid phase, the mole fraction of component 1 in the solid phase increases as $\beta$ declines. Consequently, for the system in part b of Figure 7, the composition difference between coexisting fluid and solid phases is smaller than that for the system in part a of Figure 7 when $x$ is smaller than the mole fraction of component 1 at the eutectic point $x^*$. Conversely, for the system in part b of Figure 7, the composition difference between coexisting fluid and solid phases is larger than that for the system in part a of Figure 7 when $x$ is larger than $x^*$.

For systems shown in Figure 7, phase diagrams are very sensitive to parameter $\beta$ that represents the difference in pure-component solubility. In addition, the eutectic point is also sensitive to parameter $\beta$. For $\beta=1$, 0.9, and 0.8, the eutectic composition $x^* = 0.4$, 0.16, and 0.06, respectively.

**Comparison with Experiment.** At present, for aqueous protein mixtures, there is no published work that systematically examines the variation of the fluid-solid phase diagram with experimental variables. However, available studies (Hung, 1998; Judge et al., 1995; Mahadevan and Hall, 1992; Polson, et al., 1964) indicate that crystallization from protein mixtures produces crystals of high purity when there is little affinity between unlike proteins and when the disparity in size is sufficiently large.
Upon inducing bulk crystallization with ammonium sulfate, Judge et al. (1995) demonstrated recovery and purification of ovalbumin (45 kDa) from an aqueous mixture that contains 12 g ovalbumin, 1 g conalbumin (80 kDa), and 1 g lysozyme (14.6 kDa). In their experiment, the purity of the recovered ovalbumin crystals was greater than 99%. According to the present model, the high selectivity demonstrated by this experiment may follow because ovalbumin is the major component of the protein mixture and because there is sufficient size difference between ovalbumin and conalbumin and between ovalbumin and lysozyme.

Another aqueous protein mixture wherein pure components crystallize separately is the system thaumatin (23 kDa) and lysozyme (14.6 kDa) (Hung, 1998).

Conclusions

A simple van der Waals-type theory is presented to compute phase diagrams of aqueous solutions containing binary mixtures of globular proteins. Theory is based on a one-component model that uses a composition-dependent hard-sphere diameter and a composition-dependent solvent-mediated effective temperature to represent solutions of globular proteins. Compared to other theories for protein solutions, we use a much simpler van der Waals-type perturbation theory based on the inverse-power potential of variable exponent $n$.

In a temperature-density diagram, our model correctly represents the phase diagram where the metastable fluid-fluid coexistence curve lies below the stable fluid-solid coexistence domain. For single-protein systems, despite its simplicity, the theoretical phase diagram with $n=7$ and $z=12$ shows good agreement with the observed solubility of lysozyme (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996). For our model, in terms of the effective reduced temperature $\tilde{T}$, the crystallization slot for globular proteins is approximately given by $0.18 < \tilde{T} < 0.45$.

For mixtures of globular proteins, our model takes into account the size difference between unlike proteins that has a significant effect on the partitioning of protein mixtures between coexisting fluid and solid phases. In fluid-solid phase equilibria, as the disparity in size
increases, the mutual solubility between unlike proteins declines sharply in the precipitated solid phase. For significant size differences, the solid phase approaches a pure component. The type of precipitated protein depends on the composition of the protein mixture. For fixed $\alpha$, $\beta$, and $k_{12}$, the mutual solubility in the solid phase also declines as the effective reduced temperature decreases.

For systems with large size difference, there is a eutectic point where two solid phases and a fluid phase coexist. For a given size ratio, the eutectic composition is sensitive to the solubility difference between unlike proteins.

Immiscibility in the solid phase due to the size effect declines as the interaction between unlike proteins becomes more attractive.

Acknowledgment

We thank Drs. Rosenbaum and Zukoski for providing solubility data for several colloidal solutions. This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF0098. Additional funding was provided by the Donors of the Petroleum Research Fund administered by the American Chemical Society.

Appendix

Because Rosenbaum et al. (Rosenbaum et al., 1996; Rosenbaum and Zukoski, 1996) report their data in terms of the reduced temperature for the adhesive hard-sphere potential, it is necessary to express their data in terms of the effective temperature for our model given by eq. (3). We first review the adhesive hard-sphere potential.

The adhesive hard-sphere (AHS) potential is given by
where \( \tau \) is the effective reduced temperature for the adhesive hard-sphere potential. The second virial coefficient for this potential is given by

\[
B^{AHS}_2 = 2\pi \sigma^3 \int_0^\infty \left[ 1 - e^{-\mu^{AHS}(\tilde{R}/k_B T)} \right] \tilde{R}^2 \tilde{R} = \frac{2\pi}{3} \sigma^3 \left( 1 - \frac{1}{4\tau} \right) \quad (\tilde{R} = R/\sigma) .
\]

(A 2)

For \( 0.13<\tau<2 \), the theoretical phase diagram has not been reported for the adhesive hard-sphere potential. Therefore, Rosenbaum and coworkers approximated the fluid-solid coexistence curve for the adhesive hard-sphere potential by that for the short-range attractive Yukawa potential obtained from computer simulations by Hagen and Frenkel (1994). For the short-range attractive Yukawa potential, this approximation is satisfactory when the phase diagram for the Yukawa potential is compared with that for the adhesive hard-sphere potential at the same second virial coefficient.

We express the data from Rosenbaum et al. in terms of \( \tilde{T} \) by equating the second virial coefficient for the inverse-power potential to that for the adhesive hard-sphere potential:

\[
B^{IP}_2 = B^{AHS}_2
\]

(A 3)

where \( B^{IP}_2 \) and \( B^{AHS}_2 \) are given by eqs. (2) and (A 2), respectively. Because the theoretical second virial coefficient depends on both hard-sphere diameter and effective reduced temperature, the hard-sphere diameter reported by Rosenbaum and Zukoski may be adjusted to obtain better agreement of our model with experiment. Although Rosenbaum and coworkers
obtained hard-sphere diameters such that the fluid branch of the theoretical coexistence curve agrees with experiment in the high-density region, the regressed hard-sphere diameters for several colloidal particles are very close to those estimated from other measurements.

**Literature Cited**


Figure Captions

Figure 1. Theoretical temperature-density diagrams for single-protein systems (F=fluid, S=solid): (a) \( n=5 \), (b) \( n=6 \). In the perturbation term, the solid phase is the face-centered-cubic lattice with \( z=12 \) and \( y_{cp}=0.74 \). In part b, the broken curve is the coexistence curve for the metastable fluid-fluid phase separation.

Figure 2. Phase diagram for aqueous lysozyme solution at pH 4.6. The open and solid circles are the solubility data from Rosenbaum and Zukoski (1996) for \( \sigma=3.4 \) nm and \( \sigma=3.2 \) nm, respectively. The open and solid triangles are the phase separation temperatures associated with the metastable fluid-fluid phase separation for \( \sigma=3.4 \) nm and \( \sigma=3.2 \) nm, respectively. (Data shown by open and solid symbols are from the same set of experimental data.) The solid and broken curves are theoretical fluid-solid and fluid-fluid coexistence curves, respectively, for \( n=7 \).

Figure 3. Theoretical phase diagrams for aqueous protein mixtures for \( \alpha=1 \) at \( \tilde{T}_{11}=0.45 \) (F=fluid, S=solid): (a) \( \beta\neq1 \) and \( k_{12}=0 \), (b) \( \beta=1.2 \) and \( k_{12}\neq0 \). In these figures, there is no difference in the hard-sphere diameter between components 1 and 2. For \( \beta>1 \), component 1 has larger solubility than component 2.

Figure 4. Theoretical phase diagrams for aqueous protein mixtures with different size ratios (F=fluid, S=solid): (a) \( \alpha=0.95 \), (b) \( \alpha=0.9 \), (c) \( \alpha=0.85 \). For all systems \( \beta=0, k_{12}=0 \) and \( \tilde{T}_{11}=0.45 \). In part c, there is a eutectic point where two solid phases and a fluid phase coexist.

Figure 5. Theoretical phase diagrams for aqueous protein mixtures with different \( k_{12} \) (F=fluid, S=solid): (a) \( k_{12}=-0.1 \), (b) \( k_{12}=-0.2 \). For all systems \( \alpha=0.85, \beta=1, \) and \( \tilde{T}_{11}=0.45 \).
Figure 6. Theoretical phase diagrams for aqueous protein mixtures at different effective temperatures (F=fluid, S=solid): (a) $\mathcal{T}_{11} = 0.45$, (b) $\mathcal{T}_{11} = 0.4$, (c) $\mathcal{T}_{11} = 0.35$. For all systems $\alpha = 0.9$, $\beta = 1$, and $k_{12} = 0$.

Figure 7. Theoretical phase diagrams for aqueous protein mixtures (F=fluid, S=solid): (a) $\beta = 1$, (b) $\beta = 0.9$. For all systems $\alpha = 0.9$, $k_{12} = 0$, and $\mathcal{T}_{11} = 0.35$. 
Figure 1

(a) $n=5$

$F$ $F-F$ $F-S$ $S$

(b) $n=6$

$F$ $F-S$ $S$

$k_B T/\varepsilon$ vs. $\pi \rho a^3/\delta$
Figure 2
Figure 3

(a)

\[ \frac{p \sigma_{11}^3}{\varepsilon_{11}} \]

- Dashed line: \( \alpha = 1 \), \( \beta = 1.1 \)
- Solid line: \( \alpha = 1 \), \( \beta = 1.2 \)

\( \bar{T}_{11} = 0.45 \), \( k_{12} = 0 \)

(b)

\[ \frac{p \sigma_{11}^3}{\varepsilon_{11}} \]

- Dashed line: \( k_{12} = 0.08 \)
- Solid line: \( k_{12} = -0.2 \)

\( \bar{T}_{11} = 0.45 \), \( \alpha = 1 \), \( \beta = 1.2 \)
Figure 7