Molecular Thermodynamics for Fluid-Phase Equilibria in Aqueous Two-Protein Systems

Nicolas von Solms, Camille O. Anderson, Harvey W. Blanch and John M. Prausnitz

Dept. of Chemical Engineering, University of California, Berkeley CA 94720

and

Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Abstract

With saline water as the continuous medium, a two-component McMillan-Mayer equation of state is used to describe liquid-liquid phase equilibria in a two-protein system. The equation of state is based on a hard-sphere reference with perturbations introduced through a potential of mean force to account for electrostatic forces and for attraction between protein particles. To illustrate the thermodynamic framework, one parameter each is fitted to experimental precipitation data for aqueous saline one-protein systems containing either lysozyme or ovalbumin. A lysozyme-ovalbumin interaction parameter is then introduced to calculate phase behavior in the aqueous two-protein system. These calculations are remarkably similar to classical vapor-liquid equilibrium calculations using an equation of state. For the aqueous two-protein system, calculations give the light-phase composition as well as the lysozyme and ovalbumin partition coefficients for a given dense-phase composition. Agreement with sparse experimental data is reasonable over a range of pH and high ionic strength provided by the common precipitant ammonium sulfate.

Keywords: phase equilibria, equation of state, two-protein systems, molecular thermodynamics, protein precipitation.

Topical Heading: Thermodynamics

Introduction

Precipitation from solution provides a standard practical method for protein separation (Bjurstrom, 1985). It is especially popular as the first, coarse step in a protein purification sequence. Precipitation is achieved by addition of a precipitating agent to a solution containing one or more proteins. Addition of this agent induces a phase split where the solution separates into regions that are, respectively, rich and lean in the target protein. A number of precipitation agents have been used, including organic solvents such as methanol, ethanol, isopropanol and acetone; polyelectrolytes such as polyacrylic acids, polysaccharides and polyphosphates (Hoare and Dunnill, 1984) and non-ionic polymers such as polyethylene glycol (Asakura and Oosawa, 1954, 1958; Foster et al., 1973; Vrij, 1976; de Hek and Vrij, 1981; Joanny et al., 1979; Middaugh et al., 1979; Vincent et al., 1980; Sperry et al., 1981; Atha and Ingham, 1981; Gast et al., 1983a, b; Haire et al., 1984; Forciniti et al., 1991; Mahadevan and Hall, 1992a, b). Fractionation of human blood using ethanol is perhaps the largest-volume application of protein precipitation (Cohn and Edsall, 1943; Bjurstrom, 1985). However, the most common industrial precipitation agent is an inorganic salt such as ammonium (or sodium) sulfate, citrates and phosphates (Cowel and Vincent, 1982; Grimson, 1983; Victor and Hansen, 1984; Bjurstrom, 1985; Niederauer and Glatz, 1992; Shih et al., 1992; Rothstein, 1994; Coen et al., 1995, 1997; Kuehner et al., 1996).

As biotechnology continues to expand toward large-scale production, better separation techniques for larger quantities of protein will be required. For large-scale separations, precipitation with ammonium sulfate is particularly attractive because of its low cost. Increasing volumes of protein throughput in separation processes provide motivation for a molecular thermodynamic method to aid design of protein-separation processes. While development of a reliable theory is hindered by the paucity of suitable experimental data, the work presented here reports a molecular-thermodynamic framework for a quantitative description of liquid-liquid equilibria for an aqueous system containing two or more proteins.

Most previous work on the thermodynamics of protein-precipitation has been confined to systems containing only one protein (e. g. Vlachy et al., 1993; Chiew et al., 1995; Kuehner et al., 1996; Muschol and Rosenberger, 1997; Sear, 1999; Kim et al., 2000); however, descriptions for aqueous multi-protein systems have been reported by Mahadevan and Hall (1992b) and Hino and Prausnitz (1999). These earlier studies assumed that pure, crystalline proteins are precipitated from a solution containing protein and salt (Arakawa and Timasheff, 1985). It is well known that precipitation of protein crystals requires special care and much patience. In a typical large-scale industrial application, the precipitate is not a crystal but a second fluid phase (Bjurstrom, 1985; Shih et al., 1992; Berland et al., 1992); theoretical and molecular-simulation studies suggest that the observed liquid-liquid equilibria represent metastable states (ten Wolde and Frenkel, 1997).

To represent aqueous protein phase behavior, we require a suitable potential of mean force to represent a variety of solvent-mediated protein-protein intermolecular forces.

These include (but are not limited to) electrostatic repulsion, solvent cavity formation, protein dipole-dipole forces, dispersion potentials, osmotic attraction potentials and a specific attraction potential typically represented by a square-well or Lennard-Jones attraction. Native proteins are approximately described by hard spheres with isotropic potentials (e. g. Vlachy et al., 1993; Chiew et al., 1995) but in some cases the protein is considered as a linear flexible chain or as a sphere or ellipsoid with potentials arising from "patches" on the surface of the protein molecule (Asthagiri and Lenhoff, 1997).

Kuehner et al (1996) summarized the procedure for applying an equation of state to describe liquid-liquid equilibrium in a one-protein aqueous system. For a one-protein solution in a continuous medium, the McMillan-Mayer theory is used to reduce the system to a pseudo one-component system (see, for example, McQuarrie, 1976a). McMillan and Mayer showed that the methods of imperfect-gas theory can be applied to such solutions. In this work we consider an aqueous two-protein solution. Extension to aqueous multi-protein solutions is straightforward. While, in principle the equilibria of concern her could also be described by an activity coefficient model, it is not at all clear what such a model might be because, for the calculations given here, the model must apply to a larger range of protein concentrations. The experimental data used are for dilute solutions and thus it is necessary to use some theoretical approximation for extending dilute solution data to concentrated solutions. A well-established approximation for this purpose is the random phase approximation (RPA) for establishing a perturbation to a hard-sphere reference system (Vlachy et al., 1993; Chiew et al., 1995).

Further, a necessary condition for fluid-phase equilibria is that, at the same temperature, the two phases have the same osmotic pressure; that pressure is conveniently calculated from our equation of state. In an activity-coefficient description, equality of the osmotic pressures is replaced by equality of the chemical potential of the solvent (water). Therefore, whatever model is used for activity coefficients needs a ternary model. In the activity coefficient (Lewis-Randall) description, the two-protein aqueous system is a ternary mixture. In the McMillan-Mayer description, a two-protein system is a binary mixture.

Following earlier work (e. g. Kuehner et al., 1996), we consider globular proteins to be hard spheres with a perturbation introduced via the potential of mean force. This procedure provides a large simplification because the molecular solvent is replaced by a continuous solvent medium. While the specific results of this work are only valid for the solvent (water) and salt (ammonium sulfate) choices we have studied, the thermodynamic and computational framework are general, applicable to other solvents, salts and proteins.

The potential of mean force is related to the osmotic second virial coefficient in an exact way (McQuarrie, 1976b). Because that coefficient can be measured relatively easily, we can, in principle, obtain the microscopic potential of mean force from macroscopic measurements.

Coen (1995) and Coen et al (1995, 1997) have obtained limited experimental data for the precipitation of lysozyme and ovalbumin in one-protein systems with ammonium sulfate over a range of pH and high ionic strength. Only a few data were obtained for the phase behavior of these proteins together in an aqueous two-protein system at similar pH and ionic strength. In this work we present a two-component equation of state describing liquid-liquid equilibria in aqueous two-protein systems. The specific interaction energy for lysozyme, as well as for ovalbumin (which depends on ionic strength and pH) was obtained by fitting to one-protein precipitation data of Coen (1995) and Coen et al (1995, 1997). The behavior of the two-protein system is then calculated using these single-protein interaction energies and one additional interaction parameter for the cross interaction.

The next section describes the thermodynamic framework. Sample results are then presented and discussed, leading to some conclusions.

Thermodynamic Framework

The aqueous protein solution contains a mixture of two types of hard spheres that differ in diameter and charge. The total Helmholtz energy of the system A^{total} is given by

$$\frac{A^{total}}{NkT} = \frac{A^{res}}{NkT} + \frac{A^{ideal}}{NkT} \tag{1}$$

where N is the total number of molecules in the system, k is Boltzmann's constant, T is the absolute temperature and A^{res} is the residual Helmholtz energy of the system.
The ideal Helmholtz energy A^{ideal} is given by

$$\frac{A^{ideal}}{NkT} = \sum_{i=1}^{2} x_i \ln \rho_i \Lambda_i^3 - 1 \tag{2}$$

where x_i is the mole fraction of protein *i* on a solvent-free basis, ρ_i is the number density of protein *i* and Λ_i is its de Broglie wavelength. The residual Helmholtz energy of this mixture is the sum of two parts: a reference Helmholtz energy and a perturbation Helmholtz energy. The perturbation Helmholtz energy is given by the Random Phase Approximation (RPA) (Evans and Sluckin, 1981; Grimson, 1983; Victor and Hansen, 1984; Vlachy et al., 1993; Chiew et al., 1995; Kuehner et al., 1996). The residual Helmholtz energy is the sum of two contributions:

$$\left(\frac{A^{res}}{NkT}\right) = \left(\frac{A^{res}}{NkT}\right)^{ref} + \left(\frac{A^{res}}{NkT}\right)^{pert} . \tag{3}$$

The reference Helmholtz energy is given by the BMCSL equation of state (Mansoori et al., 1971; Lucas, 1991). The compressibility factor Z^{ref} is

$$Z^{ref} = \frac{P^{ref}}{\rho kT} = \frac{1}{1-\xi_3} + \frac{3\xi_1\xi_2}{\xi_0(1-\xi_3)^2} + \frac{\xi_2^3(3-\xi_3)}{\xi_0(1-\xi_3)^3} \tag{4}$$

where ρ is the total protein number density. The quantities ξ_k are measures of composition defined by

$$\xi_k = \frac{\pi}{6}\sum_i \rho_i \sigma_i^k , \qquad k = 0,1,2,3 \tag{5}$$

where index *i* refers to the protein and σ_i is the hard-sphere diameter of protein *i*. In Eq. 5 subscript and superscript *k* should not be confused with subscript *i*. When *k* = 3, Eq. 5 gives the reduced protein density η.

The residual Helmholtz energy is obtained from the compressibility factor using the thermodynamic relation

$$\left(\frac{A^{res}}{NkT}\right)^{ref} = \int_0^{\eta} (Z^{ref} - 1)\frac{d\eta}{\eta} \tag{6}$$

to yield

$$\left(\frac{A^{res}}{NkT}\right)^{ref} = \frac{\xi_2^3 + 3\xi_1\xi_2\xi_3(1-\xi_3)}{\xi_0\xi_3(1-\xi_3)^2} - \left(1 - \frac{\xi_2^3}{\xi_0\xi_3}\right)\ln(1-\xi_3). \tag{7}$$

The chemical potential of component *i* in the mixture is obtained using the thermodynamic relation

$$\left(\mu_i^{res}\right)^{ref} = \left(\frac{\partial A^{res}}{\partial N_i}\right)^{ref}_{T,V,N_{j\neq i}} . \tag{8}$$

After lengthy algebra we obtain

$$\left(\frac{\mu_i^{res}}{kT}\right)^{ref} = \left[\frac{\xi_2^3}{\xi_3^2}\left(\frac{3\sigma_i}{\xi_1} - \frac{2\sigma_i^3}{\xi_3}\right) - 1\right] \ln(1-\xi_3)$$
$$+\frac{1}{1-\xi_3}\left[\left(1-\frac{\xi_2^3}{\xi_3^2\xi_0}\right)\sigma_i^3\xi_0 + 3\xi_1\xi_2\left(\frac{\sigma_i}{\xi_1} + \frac{\sigma_i^2}{\xi_2} - \frac{\sigma_i^3}{\xi_3}\right)\right] \tag{9}$$
$$+\frac{1}{(1-\xi_3)^2}\left[\frac{3\xi_1\xi_2\sigma_i^3}{\xi_3} + \frac{3\xi_2^3}{\xi_3}\left(\frac{\sigma_i^2}{\xi_2} - \frac{\sigma_i^3}{\xi_3}\right)\right] + \frac{2\xi_2^3\sigma_i^3}{\xi_3^2(1-\xi_3)^3}$$

The reference chemical potential contains all contributions other than those introduced by the perturbation; it is the sum of two parts, the residual chemical potential (Eq. 9) and the contribution of the ideal portion obtained from Eq. 2,

$$\left(\frac{\mu_i}{kT}\right)^{ref} = \left(\frac{\mu_i^{res}}{kT}\right)^{ref} + \frac{\mu_i^{id}}{kT}. \tag{10}$$

The ideal chemical potential is:

$$\left(\frac{\mu_i^{id}}{kT}\right) = \ln\rho_i + \ln\Lambda_i^3. \tag{11}$$

Because Λ_i is independent of density and composition, it cancels out in phase-equilibrium calculations at constant temperature.

The Random Phase Approximation (RPA) gives a simple expression for the contribution form perturbation forces to the compressibility factor:

$$Z^{pert} = \frac{\rho U}{2kT} \tag{12}$$

where U is related to the perturbation potential of mean force $w^{pert}(r)$ and to composition x_i, the mole fraction of component i on a solvent-free basis:

$$U = 4\pi \sum_i \sum_j x_i x_j \int_{\sigma_{ij}}^{\infty} w_{ij}^{pert}(r;\sigma_{ij}) r^2 dr\,. \tag{13}$$

For proteins in salt solutions, the potential of mean force contains three contributions arising from dispersion, electrostatic and specific attractive forces. While these three contributions do not necessarily include all possible perturbation interactions, previous experience (Kuehner et al., 1996) suggests that they adequately describe the observed behavior in systems such as ours. As shown later, the contribution arising from specific interactions often far outweigh the other two contributions. We have:

$$w_{ij}^{pert}(r) = w_{ij}^{disp}(r) + w_{ij}^{elec}(r) + w_{ij}^{spec}(r) \tag{14}$$

where

$$
\begin{aligned}
w_{ij}^{disp}(r) &= -\frac{H}{\left(\frac{r}{\sigma_{ij}}\right)^{6}} \\
w_{ij}^{elec}(r) &= \frac{z_i z_j e^2 \Gamma(r;\sigma_{ij})}{Dr} \\
w_{ij}^{spec} &= \left.\begin{aligned} &-\varepsilon_{ij};\sigma_{ij} < r < \lambda_{ij}\sigma_{ij} \\ &= 0; r > \lambda_{ij}\sigma_{ij} \end{aligned}\right\}
\end{aligned}
\qquad (15)
$$

The dispersion term is essentially identical to the attractive part of the Lennard-Jones potential scaled by *H*, the Hamaker constant, that varies little from one globular protein to another. As suggested by Hiemenz (1986), we use $H = 5\ kT$. For w_{ij}^{spec} we use a square-well potential where ε_{ij} is the well depth and $\lambda_{ij} - 1$ is the reduced well width. This is the simplest potential that incorporates short-range attraction beyond the protein's hard-core diameter.

When $i = j$, $w_{11}^{pert}(r)$ refers to ovalbumin-ovalbumin interactions and $w_{22}^{pert}(r)$ refers to lysozyme-lysozyme interactions. When $i \neq j$, $w_{12}^{pert}(r)$ refers to the interaction between one molecule of ovalbumin and one molecule of lysozyme. To determine $w_{12}^{pert}(r)$, we use the standard combining relations

$$
\begin{aligned}
\sigma_{12} &= \frac{\sigma_{11} + \sigma_{22}}{2} \\
\varepsilon_{12} &= \sqrt{\varepsilon_{11}\varepsilon_{22}}\,(1 - k_{12}) \\
\lambda_{12} &= \frac{\sigma_{11}\lambda_{11} + \sigma_{22}\lambda_{22}}{\sigma_{11} + \sigma_{22}}
\end{aligned}
\qquad (16)
$$

where k_{12} is the only adjustable binary parameter. Because this parameter has a marked effect on phase equilibria, it is adjusted to give the best fit to the limited two-protein experimental data (Coen et al., 1995, 1997; Coen, 1995). For a fixed temperature, pH and ionic strength, the same k_{ij} is used to predict partition coefficients for both proteins, as well as the light-phase solvent-free mole fraction for a given dense-phase mole fraction. In general, in multi-protein systems, interactions between like proteins may be very different from those between unlike proteins (because different sites may be involved). The use of mixing rules might then be inappropriate. However, in the context of our simple model, and for the system here under consideration, the conventional mixing and combining rules above are appropriate and sufficiently flexible due to our binary interaction parameter k_{ij}.

The electrostatic potential of mean force is given by a screened Coulombic potential where z_i is the protein valence and e is the charge on an electron [1.519 $(\text{JÅ})^{1/2}$]. The product $z_i e$ is the charge on protein molecule i and D is the permittivity of the solvent. From Debye-Hückel theory (McQuarrie, 1976c), the screening factor $\Gamma(r;\sigma_{ij})$ is given by

$$\Gamma(r,\sigma_{ij}) = \frac{e^{-\kappa(r-\sigma_{ij})}}{(1+\frac{\kappa\sigma_{ij}}{2})^2} \tag{17}$$

where κ is the inverse Debye length given by

$$\kappa = \left(\frac{8\pi e^2 I N_A \times 10^{-27}}{DkT} \right)^{1/2} . \tag{18}$$

In Eq. 18, N_A is Avogadro's number and the factor $N_A \times 10^{-27}$ converts ionic strength (I) to units of molecules/Å^3. In Eq. 18, I is in molar units. The Debye length κ^{-1} is then in units of Å. In typical cases, I refers only to the added salt because the contribution of the proteins to I is negligible. The protein charge z_i is determined experimentally from titration curves; z_i depends strongly on pH and weakly on molarity and type of salt (Kuehner et al., 1999; Cohn and Edsall, 1943). For the high ionic strengths considered here, the effect of electrostatic interactions on phase equilibria is very small. Table 1 gives z_i for ovalbumin and lysozyme at different pH.

Square-well parameters λ_{ij} and ε_{ij} in Eq. 15 are often obtained from experimental osmotic-second-virial-coefficient data or from cloud-point data (Grigsby, 2001). As suggested by osmotic-second-virial-coefficient data (Gast et al., 1983b, 1986; George and Wilson, 1994), the reduced square-well width λ is set equal to 1.2 for both lysozyme and ovalbumin, consistent with the molecular-simulation data of ten Wolde and Frenkel (1997). Following Eq. 16, the cross-parameter λ_{12} is also equal to 1.2. The square-well depth parameters ε_{11} and ε_{22} for the two one-protein systems are obtained by fitting the partition coefficients from the equation of state to those obtained from experiment (Coen, 1995b) for the one-protein systems. At the high ionic strengths considered here, $w^{spec}(r)$ provides the dominant contribution to $w^{pert}(r)$.

For the potential of mean force used here, we have three integrals, J_{ij} in Eq. 13. They can be solved analytically:

$$\begin{aligned} J_{ij}^{disp} &= \int_{\sigma_{ij}}^{\infty} w_{ij}^{disp} r^2 dr = -\frac{H\sigma_{ij}^3}{108} \\ J_{ij}^{elec} &= \int_{\sigma_{ij}}^{\infty} w_{ij}^{elec} r^2 dr = \frac{z_i z_j e^2 (\sigma_{ij} + \frac{1}{\kappa})}{D(1+\frac{\kappa\sigma_{ij}}{2})^2 \kappa} \\ J_{ij}^{spec} &= -\frac{\varepsilon_{ij}\sigma_{ij}^3}{3}(\lambda_{ij}^3 - 1) \end{aligned} \qquad (19)$$

The perturbation Helmholtz energy is obtained by substituting Eq. 12 into the thermodynamic relation Eq. 6:

$$\frac{A^{pert}}{NkT} = \frac{\rho U}{2kT} = \frac{2\pi\rho}{kT}\sum_i \sum_j x_i x_j J_{ij}^{total} \qquad (20)$$

We can now find the perturbation contribution to the chemical potential:

$$\left(\frac{\mu_i}{kT}\right)^{pert} = \frac{\partial}{\partial N_i}\left(\frac{A}{kT}\right)^{pert} = \frac{4\pi\rho}{kT}\sum_j x_j J_{ij}^{total} \qquad (21)$$

where J_{ij}^{total} is the sum of the integrals in Eq. 19. We now have all the expressions required to compute phase equilibria from

$$\mu_i' = \mu_i'' \quad , \qquad i = 1,2 \tag{22}$$

where superscripts $'$ and $''$ stand, respectively, for the dense-fluid and light-fluid phases and from

$$P' = P'' \tag{23}$$

where P is pressure. The phase diagrams are calculated by substituting Eqs. 9, 11 and 21 into Eq. 22 for the chemical potential, substituting Eqs. 4 and 12 into Eq. 23 for the pressure, and solving these three equations simultaneously. The dense-phase composition x_i' is fixed for each phase equilibrium calculation. To generate complete phase diagrams, values of x_1' (ovalbumin dense-phase mole fraction) ranging from 0 to 1 are chosen. The initial calculation is done for $x_1' = 0$ (i.e. a one-protein lysozyme system). For each subsequent calculation, x_1' is increased by 0.05 and the calculation repeated until $x_1' = 1$, a one-protein ovalbumin system. This increment (0.05) is small enough to generate smooth curves on the phase diagrams. For most of the two-protein systems considered here, phase separation occurs over the whole range of x_i'. If, however, a value is chosen for which no phase separation occurs, the equilibrium conditions Eq. 22 and 23 can not be met for dense-fluid and light-fluid densities which are different from each other. We are then in a one-phase region. As we approach the limit of the two-phase region, the partition coefficients for both proteins tend to unity. The three unknowns (overall phase densities ρ^{dp}, ρ^{lp} and light-phase mole fraction x_i'') are determined by simultaneously solving the three equations of equilibrium, Eqs. 22 and 23. This calculation is identical to a bubble-point calculation in vapor-liquid equilibria. The equations are solved iteratively using Broyden's method (Press et al., 1992) for finding the simultaneous numerical solutions to Eqs. 22 and 23.

Thermodynamic calculations can be compared with experimental data in the following way: because solvent (with dissolved salt) is considered to be a continuous medium, a one-protein system is a pure-component system in a solvent-free framework. Experimental partition coefficients reported by Coen (1995) and Coen et al. (1995, 1997) for a one-protein system are given in terms of concentrations:

$$K^{one-protein} = \frac{c^{dp}}{c^{lp}} \tag{24}$$

where c^{dp} is the protein concentration in the dense phase in mg/g water and c^{lp} is the protein concentration in the supernatant (light) phase. From the equation of state for a one-protein system we obtain the molar densities of the light phase and dense phase. The partition coefficient can be written

$$K^{one-protein} = \frac{\rho^{dp}}{\rho^{lp}} \tag{25}$$

where ρ_{dp} is the dense-phase protein number density and ρ_{lp} is the light-phase protein number density.

In the two-protein system, consistent with Eq. 24, we have two distribution coefficients

$$K_1^{two-protein} = \frac{c_1^{dp}}{c_1^{lp}} \quad \text{and} \quad K_2^{two-protein} = \frac{c_2^{dp}}{c_2^{lp}}. \tag{26}$$

Because water and dissolved salt are a continuous medium they contribute no mass to the calculated protein densities. Eq. 26 can be rewritten

$$K_1^{two-protein} = \frac{x_1^{'} \rho_{dp}}{x_1^{''} \rho_{lp}} \quad \text{and} \quad K_2^{two-protein} = \frac{x_2^{'} \rho_{dp}}{x_2^{''} \rho_{lp}} \tag{27}$$

where $x_i^{'}$ is the solvent-free mole fraction of component *i* in the dense phase and $x_i^{''}$ is the solvent-free mole fraction in the light phase. For comparison with experiment, concentration data reported in mass units are converted to solvent-free mole fractions by:

$$x_i = \frac{\dfrac{c_i}{\left(MW\right)_i}}{\sum_{j=1}^{2} \dfrac{c_j}{\left(MW\right)_j}} \tag{28}$$

where x_i is a solvent-free mol fraction in either phase, c_i is a mass concentration in that phase and $\left(MW\right)_i$ is the molecular weight of component *i*: 14 300 g/mol for lysozyme and 45 000 g/mol for ovalbumin.

Results and Discussion

For data reduction, we need molecular diameters for both proteins. From crystallographic data we use 50 Å for ovalbumin (Stein et al., 1990) and 34.4 Å for lysozyme (Blake et al., 1965). For each protein, well depth ε depends on pH and on ionic strength of ammonium sulfate solution, as shown in Tables 2 and 3. The well depths were obtained from experimental one-protein liquid-liquid equilibrium data (Coen, 1995; Coen et al., 1995, 1997) at 25 °C.

For the two-protein system, we require binary parameter k_{12} as defined in Eq. 16. This binary parameter also depends on pH and ionic strength of ammonium sulfate as shown in Table 4.

The molecular-thermodynamic framework presented here enables us to calculate phase equilibria in a manner virtually identical to that for vapor-liquid equilibria as used by chemical engineers for about fifty years. In both cases, the calculation requires an equation of state and characteristic constants that reflect intermolecular forces: for a binary mixture containing components 1 and 2, we need one set of constants for 1-1 interactions, another for 2-2 interactions and another for 1-2 interactions. The first and second sets we obtain from experimental data for the one-component systems but the third set requires some binary data. Thus, there is a striking similarity between conventional calculations for vapor-liquid equilibria and those for aqueous multi-protein systems. Regrettably, the latter require more experimental information because intermolecular forces between proteins, unlike those for conventional non-electrolyte fluids, depend on temperature, pH, ionic strength and the nature of the salt in the aqueous medium.

Figures 1 to 5 show some calculated results for the aqueous two-protein system ovalbumin - lysozyme. For a fixed pH and ionic strength, each figure shows a binary $x^{'} - x^{''}$ diagram as well as the corresponding two-protein partition coefficients for each protein. Figures 1 to 3 show results for pH 7, Figures 4 and 5 for pH 4. Figure 1 shows phase behavior at pH 7 and ionic strength 9 molal ammonium sulfate. The binary phase diagram 1(a) is typical of those observed in vapor-liquid systems; no unusual behavior is observed. The calculated *K*-values 1(b) are in reasonable agreement with experiment, especially considering that the accuracy of the data is not high. Figure 2 at pH 7 and somewhat lower ionic strength 8 molal shows azeotropic behavior. Figure 3 at pH 7 and ionic strength 6 molal indicates phase separation only up to a solvent-free mole fraction of ovalbumin of about 0.4; at higher ovalbumin concentrations there is no precipitation. As the mole fraction of ovalbumin increases up to 0.4, the partition coefficients decrease to unity. At higher ovalbumin mole fractions no phase separation occurs. Figure 4(a) at pH 4 and ionic strength 7 molal shows azeotropic behavior similar to that in Figure 2(a) except that the deviation from ideality is larger. Finally, Figure 5 shows a highly non-ideal system at pH 4 and ionic strength 8 molal. The interaction parameter (k_{12}) here is the largest of any of the systems studied. The calculations indicate a region where a third fluid phase appears when the ovalbumin solvent-free mole fraction in the dense phase is in the range 0.05 to 0.4.

Conclusions

A molecular-based equation of state, suitable for aqueous globular proteins, is extended to two-protein systems. This equation correctly predicts liquid-liquid phase separation when ammonium sulfate at high ionic strength is added to an aqueous mixture of ovalbumin and lysozyme. The equation is based on a hard-sphere mixture equation of state with perturbations from several contributions to the potential of mean force. A single specific energy parameter is adjusted to model phase separation in one-protein systems. Using these parameters and one binary specific energy parameter, calculated phase behavior gives reasonable agreement with sparse experimental data. Depending on pH and ionic strength, the calculations predict different types of phase behavior that are remarkably similar to those obtained from classical equation-of-state vapor-liquid-equilibrium calculations for non-electrolytes.

Acknowledgments

For financial support, the authors are grateful to the Office for Basic Science, U.S. Department of Energy, and to the National Science Foundation.

Literature Cited

Arakawa, T. and S. N. Timasheff, "Theory of Protein Solubility," *Methods Enzymol.,* **114,** 49 (1985).

Asakura, S., and F. Oosawa, "On Interaction Between Particles Suspended in Solutions of Macromolecules," *J. Chem. Phys.,* **22,** 1255 (1954).

Asakura, S., and F. Oosawa, "Interaction Between Particles Suspended in Solutions of Macromolecules," *J. Polym. Sci.,* **33,** 183 (1958).

Asthagiri D, and A. M. Lenhoff, "Influence of structural details in modeling electrostatically driven protein adsorption," *Langmuir*, **13**, 6761 (1997)

Atha, D. K., and K. C. Ingham, "Mechanism of Precipitation of Proteins by Polyethylene Glycols," *J. Biol. Chem.,* **256,** 12108 (1981).

Berland, C. R, G. M. Thurston, M. Kondo, M. L. Broide, J. Pande, O. Ogun, and G. B. Benedek, "Solid Liquid-Phase Boundaries of Lens Protein Solutions," *P. Natl. Acad. Sci. USA,* **89,** 1214 (1992).

Bjurstrom, E., "Biotechnology: Fermentation and Downstream Processing," *Chem. Eng.,* **92(4),** 126 (1985).

Blake, C. C. F., D. F. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips, and V. R. Sarma, "Structure of Hen Egg-White Lysozyme. A 3-Dimensional Fourier Synthesis At 2A Resolution," *Nature,* **206,** 757 (1965).

Chiew, Y. C., D. E. Kuehner, H. W. Blanch, and J. M. Prausnitz, "Molecular Thermodynamics for Salt-Induced Protein Precipitation," *AIChE J.*, **41**, 2150 (1995).

Coen, C. J., H. W. Blanch, and J. M. Prausnitz, “Salting Out of Aqueous Proteins: Phase Equilibria and Intermolecular Potentials,” *AIChE J.,* **41,** 996 (1995).

Coen, C. J. Ph.D. Thesis, University of California, Berkeley (1995).

Coen, C. J., J. M. Prausnitz, and H. W. Blanch, “Protein Salting-Out: Phase Equilibria in Two-Protein Systems,” *Biotechnol. Bioeng.,* **53,** 567 (1997).

Cohn, E. J., and J. T. Edsall, *Proteins, Amino Acids and Peptides,* Reinhold, New York (1943).

Cowel, C., and B. Vincent, “Temperature-Particle Concentration Phase Diagram for Dispersions of Weakly Interacting Particles,” *J. Colloid. Interf. Sci.,* **87,** 518 (1982)

de Hek, H., and A. Vrij, “Interactions in Mixtures of Colloidal Silica Spheres and Polystyrene Molecules in Cyclohexane,” *J. Colloid. Interf. Sci.,* **84,** 409 (1981).

Evans, R., and T. J. Sluckin, “The Role of Attractive Forces in the Structure of Simple Liquids: A Theory for Small-Angle Scattering,” *J. Phys. C,* **14,** 2569 (1981).

Forciniti, D., C. K. Hall, and M. R. Kula, “Protein Partitioning at the Isoelectric Point: Influence of Polymer Molecular Weight and Concentration and Protein Size,” *Biotechnol. & Bioeng.,* **38,** 986 (1991).

Foster, P. R., P. Dunhill, and M. D. Lilly, “The Precipitation of Enzymes from Cell Extracts of Saccharomyces Cerevisiae by Polyethylene-Glycol,” *Biochem. Biophys. Acta,* **317,** 505 (1973).

Gast, A. P., C. K. Hall, and W. B. Russel, “Phase Separations Induced in Aqueous Colloidal Suspensions by Dissolved Polymer,” *Farad. Discus. Chem. Soc.,* **76,** 189 (1983a).

Gast, A. P., C. K. Hall, and W. B. Russel, “Polymer-Induced Phase Separations in Nonaqueous Colloidal Suspensions,” *J. Colloid. Interf. Sci.,* **96,** 251 (1983b).

Gast, A. P., W. B. Russel, and C. K. Hall, “An Experimental and Theoretical-Study of Phase-Transitions in the Polystyrene Latex and Hydroxyethylcellulose System,” *J. Colloid. Interf. Sci.* **109,** 161 (1986).

George, A., and W. W. Wilson, “Predicting Protein Crystallization from a Dilute-Solution Property,” *Acta Crystallogr. D,* **50,** 361 (1994).

Grigsby, J. Ph.D. Thesis, University of California, Berkeley (2001).

Grimson, M. J., “Small-Angle Scattering from Colloidal Dispersions,” *J. Chem. Soc., Farad. Trans.,* **79(2),** 817 (1983).

Haire R. N., W. A. Tisel, J. G. White, and A. Rosenberg, “On the Precipitation of Proteins by Polymers: The Hemoglobin-Polyethylene Glycol System,” *Biopolym.,* **23,** 2761 (1984).

Hiemenz, P. C., *Principles of Colloid and Surface Chemistry,* Marcel Dekker, New York (1986).

Hino, T. and J. M. Prausnitz, “Solid-Liquid Equilibria for Solutions of Binary Globular-Protein Mixtures,” *AIChE J.,* **45,** 622 (1999).

Hoare, M., and P. Dunnill, “Improvement in Process Design for Protein Purification by Precipitation,” presented at the conference of the Engineering Foundation (New York), Recovery of Fermentation Product, Sea Island, GA (1984).

Joanny, J. F., L. Leibler, and P. G. de Gennes, “Effects of Polymer Solutions on Colloid Stability,” *J. Polym. Sci. Polym. Phys. Ed.,* **17,** 1073 (1979).

Kim, S. G., Y. C. Bae, and S. O. Ryu, “Salt-Induced Protein Precipitation in Aqueous Solution: The Effect of Pre-Aggregation,” *Kor. J. Chem. Eng.*, **17**, 638 (2000).

Kuehner, D. E., H. W. Blanch and J. M. Prausnitz, “Salt-Induced Protein Precipitation: Phase Equilibria from an Equation of State,” *Fluid Phase Equil.,* **116,** 140 (1996).

Kuehner, D. E., J. Engmann, F. Fergg, M. Wernick, H. W. Blanch and J. M. Prausnitz, “Lysozyme Net Charge and Ion Binding in Concentrated Aqueous Electrolyte Solutions,” *J. Phys. Chem. B,* **108,** 1368 (1999).

Lucas, K., *Applied Statistical Thermodynamics,* Springer-Verlag, Berlin (1991).

Mahadevan, H. and C. K. Hall, “Experimental Analysis of Protein Precipitation by Polyethylene Glycol and Comparison with Theory,” *Fluid Phase Equil.,* **78,** 297 (1992a).

Mahadevan, H. and C. K. Hall, “Theory of Precipitation of Protein Mixtures by Nonionic Polymer,” *AIChE J.,* **38,** 573 (1992b).

Mansoori, G. A., N. F. Carnahan, K. E. Starling and T. W. Leland, “Equilibrium Thermodynamic Properties of the Mixture of Hard Spheres,” *J. Chem. Phys.,* **54,** 1523 (1971).

McQuarrie, D. A., *Statistical Mechanics,* HarperCollins, New York, p. 326 (1976a).

McQuarrie, D. A., *Statistical Mechanics,* HarperCollins, New York, p. 235 (1976b).

McQuarrie, D. A., *Statistical Mechanics,* HarperCollins, New York, p. 335 (1976c).

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in Fortran,* 2nd ed. Cambridge University Press, Cambridge (1992).

Middaugh, C. R., W. A. Tisel, R. N. Haire, and A. Rosenberg, "Determination of the Apparent Thermodynamic Activities of Saturated Protein Solution," *J. Biol. Chem.,* **254,** 367 (1979).

Muschol, M., and F. Rosenberger, "Liquid-Liquid Phase Separation in Supersaturated Lysozyme Solutions and Associated Precipitate Formation/Crystallization," *J. Chem. Phys.,* **107,** 1953 (1997).

Niederauer, M. Q., and C. E. Glatz, "Selective Precipitation," *Adv. Biochem. Eng. Technol.,* **47,** 159 (1992).

Rothstein, F., "Differential Precipitation of Proteins," in *Protein Purification Process Engineering,* R. G. Harrison, ed., Marcel Dekker, New York, p. 115 (1994).

Sear, R. P., "Phase Behavior of a Simple Model of Globular Proteins," *J. Chem. Phys.,* **111,** 4800 (1999).

Shih, Y.-C., J. M. Prausnitz and H. W. Blanch, "Some Characteristics of Protein Precipitation by Salts," *Biotech. Bioeng.,* **40,** 1155 (1992).

Sperry, P. R., H. B. Hopfenberg, and N. L. Thomas, "Flocculation of Latex by Water-Soluble Polymers: Experimental Confirmation of a Nonbridging, Nonadsorptive, Volume-Restriction Mechanism," *J. Colloid. Interf. Sci.,* **82,** 62 (1981).

Stein, P. E., A. G. W. Leslie, J. T. Finch, W. G. Turnell, P. J. McLaughlin, and R. W. Carrell, "Crystal Structure of Ovalbumin as a Model for the Reactive Centre of Serpins," *Nature,* **347,** 99 (1990).

ten Wolde, P. R., and Frenkel, D., "Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations," *Science,* **277,** 1975 (1997).

Verwey, E. J. W., and J. T. K. Overbeek, *Theory of Stability of Lyophobic Colloids,* Elsevier, Amsterdam (1948).

Victor, J. M., and J.-P. Hansen, "Liquid-Gas Transition in Charged Colloidal Dispersions," *J. Physique Lett.,* **45,** L-307 (1984).

Vincent, B., P F. Luckham, and F. A. Waite, "The Effect of Free Polymer on the Stability of Sterically Stabilized Dispersions," *J. Colloid. Interf. Sci.,* **73,** 508 (1980).

Vlachy, V., H. W. Blanch, and J. M. Prausnitz, "Liquid-Liquid Phase Separations in Aqueous Solutions of Globular Proteins," *AIChE J.*, **39,** 215 (1993).

Vrij, A., "Polymers at Interfaces and the Interactions in Colloidal Dispersions," *Pure & Appl. Chem.,* **48,** 471 (1971).

Figure Captions

Figure 1. Phase diagrams for the system ovalbumin(1)-lysozyme(2) at pH 7 and ionic strength 9 molal ammonium sulfate at 25 °C. (a) Binary x'-x'' diagram showing solvent-free mole fraction in the light (double prime) and dense (prime) aqueous phases. (b) Partition coefficients. The points are from experimental solubility data of Coen (1995) and Coen et al (1995). The interaction parameter k_{12} is defined through the equation $\varepsilon_{12} = (\varepsilon_{11}\varepsilon_{22})^{1/2}(1 - k_{12})$.

Figure 2. Phase diagrams for the system ovalbumin(1)-lysozyme(2) at pH 7 and ionic strength 8 molal ammonium sulfate at 25 °C. (a) Binary x'-x'' diagram showing solvent-free mole fraction in the light (double prime) and dense (prime) aqueous phases. (b) Partition coefficients. The points are from experimental solubility data of Coen (1995) and Coen et al (1995). The interaction parameter k_{12} is defined through the equation $\varepsilon_{12} = (\varepsilon_{11}\varepsilon_{22})^{1/2}(1 - k_{12})$.

Figure 3. Partition coefficients for the system ovalbumin(1)-lysozyme(2) at pH 7 and ionic strength 6 molal ammonium sulfate at 25 °C. The points are from experimental solubility data of Coen (1995) and Coen et al (1995). The interaction parameter k_{12} is defined through the equation $\varepsilon_{12} = (\varepsilon_{11}\varepsilon_{22})^{1/2}(1 - k_{12})$.

Figure 4. Phase diagrams for the system ovalbumin(1)-lysozyme(2) at pH 4 and ionic strength 7 molal ammonium sulfate at 25 °C. (a) Binary x'-x'' diagram showing solvent-free mole fraction in the light (double prime) and dense (prime) aqueous phases. (b) Partition coefficients. The points are from experimental solubility data of Coen (1995) and Coen et al (1995). The interaction parameter k_{12} is defined through the equation $\varepsilon_{12} = (\varepsilon_{11}\varepsilon_{22})^{1/2}(1 - k_{12})$.

Figure 5. Phase diagrams for the system ovalbumin(1)-lysozyme(2) at pH 4 and ionic strength 8 molal ammonium sulfate at 25 °C. (a) Binary x'-x'' diagram showing solvent-free mole fraction in the light (double prime) and dense (prime) aqueous phases. (b) Partition coefficients. The points are from experimental solubility data of Coen (1995) and Coen et al (1995). The interaction parameter k_{12} is defined through the equation $\varepsilon_{12} = (\varepsilon_{11}\varepsilon_{22})^{1/2}(1 - k_{12})$.

Table 1. Net charge z_i for ovalbumin and lysozyme in 1 M potassium chloride for different pH. Data are from Cohn and Edsall (1943) for ovalbumin and from Kuehner et al (1994) for lysozyme.

charge vs. pH	**4**	**5**	**6**	**7**	**8**	**9**
ovalbumin	20	-2	-8	-12	-14	-16
lysozyme	14	11	9	8	7.5	7

Table 2. Specific energy parameter (square-well depth) ε / kT for ovalbumin in an aqueous one-protein system containing ammonium sulfate at 25 °C. Ionic strength I is in molal units. Values for pH 7 and ionic strength 6 and 7 molal were extrapolated from experimental data at higher ionic strengths since the ovalbumin one-protein system does not phase-separate at these conditions.

I \ pH	3	4	5	6	7
6					2.62
7	4.82	5.41	4.77		3.46
8	5.56	6.07	5.63	4.79	4.3
9	6.33	6.94	6.22	5.29	5.05
10	6.9	7.13	7.01	6.25	6.07
11			7.49	6.8	6.74
12			8.22	6.95	7.01

Table 3. Specific energy parameter (square-well depth) ε / kT for lysozyme in an aqueous one-protein system containing ammonium sulfate at 25 °C. Ionic strength I is in molal units.

I \ pH	**3**	**4**	**5**	**6**	**7**	**8**
5		5.56	3.94	3.95	3.96	3.94
6		5.7	4.52	4.27	4.33	4.34
7	6.99	6.09	5.22	4.85	5.1	5.03
8	7.25	6.33	5.85	5.42	5.57	5.57
9	7.32	7.01	6.47	6.24	6.39	6.34

Table 4. Interaction parameters k_{ij} for the aqueous two-protein system containing ammonium sulfate at 25 °C. Ionic strength I is in molal units.

I \ pH	**4**	**5**	**6**	**7**
6				-0.051
7	0.06			0.01
8	0.1			0.04
9	-0.075	-0.055	0.003	-0.02

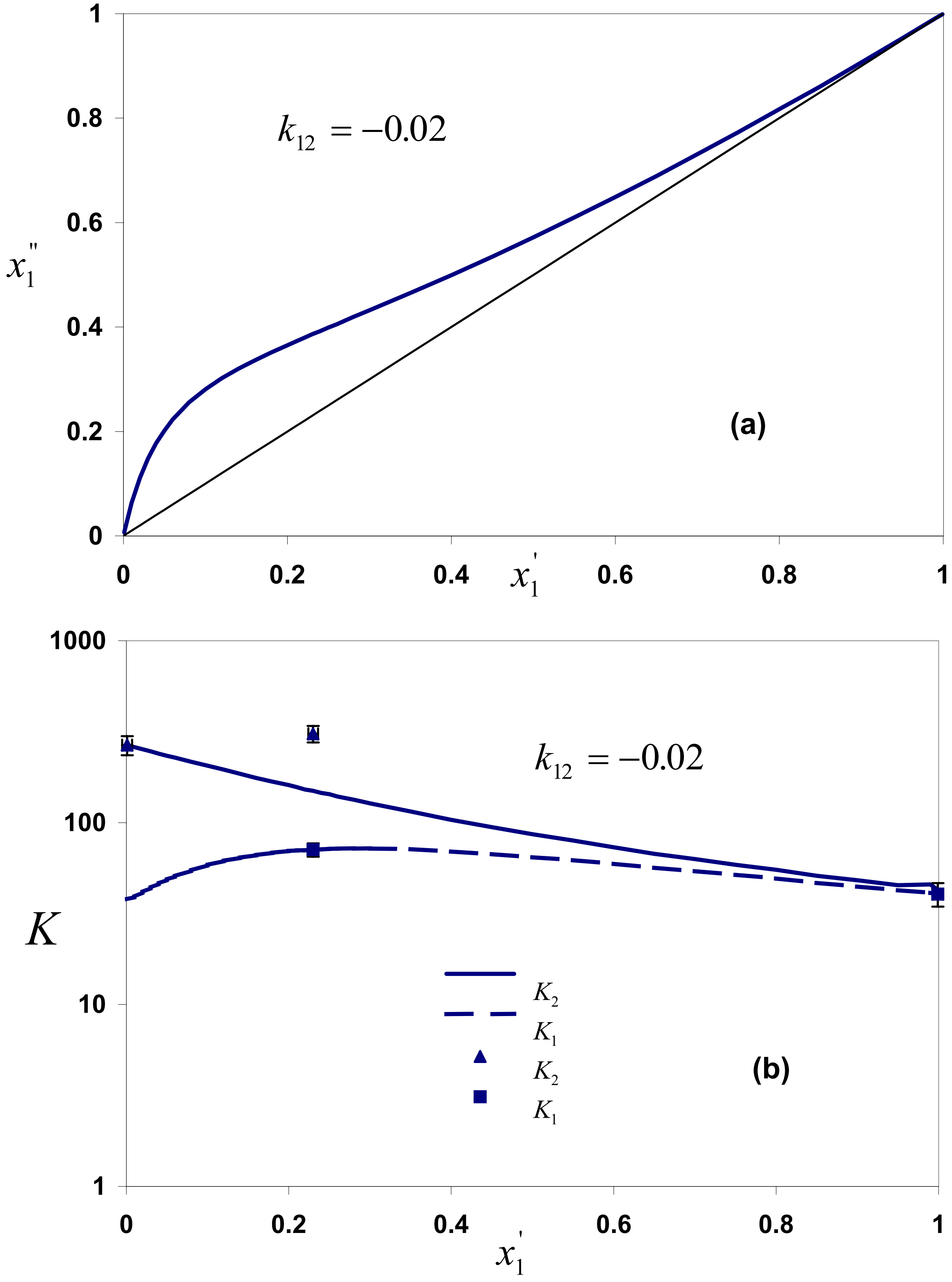

Figure 1 von Solms et al

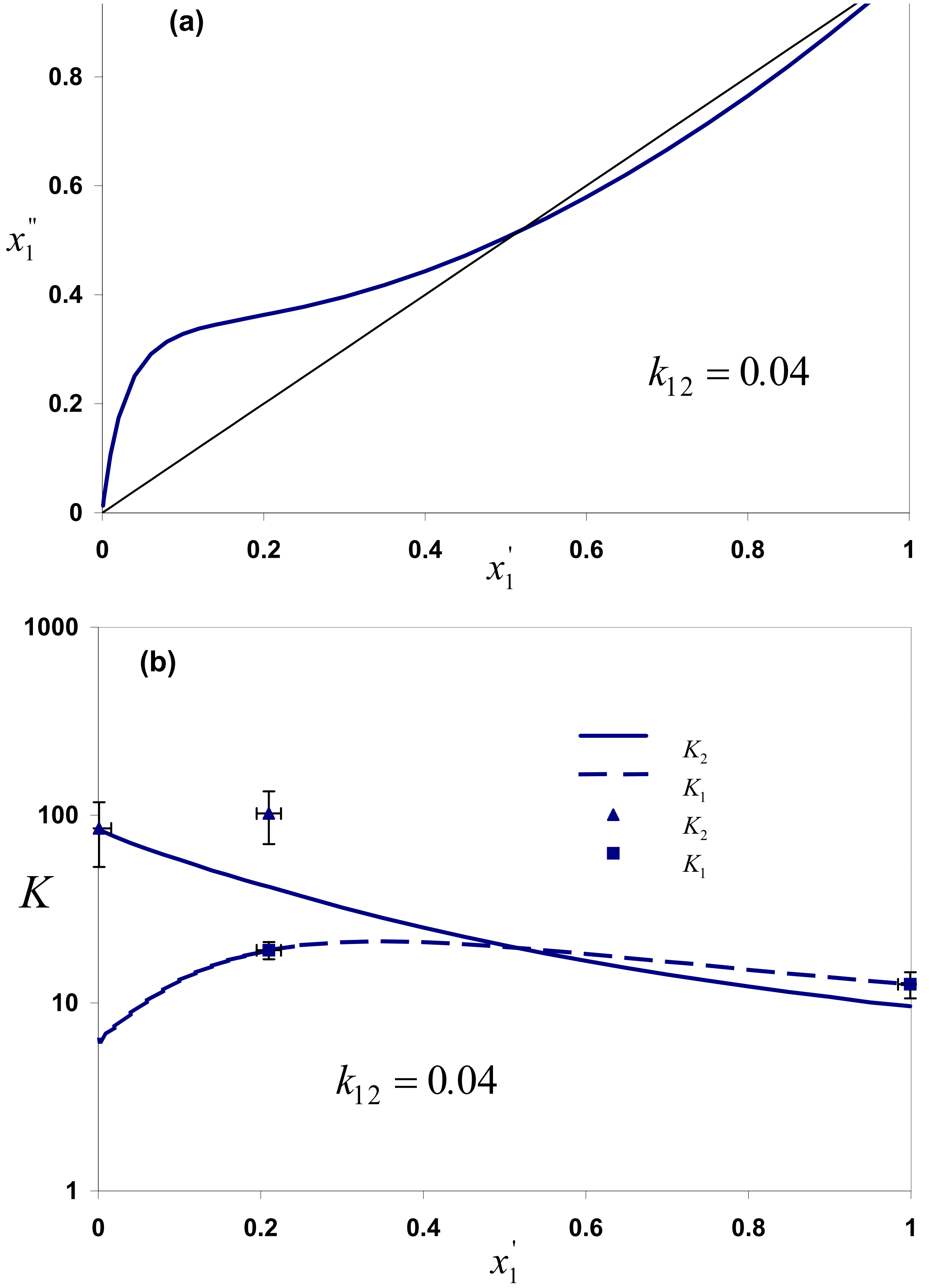

Figure 2 von Solms et al

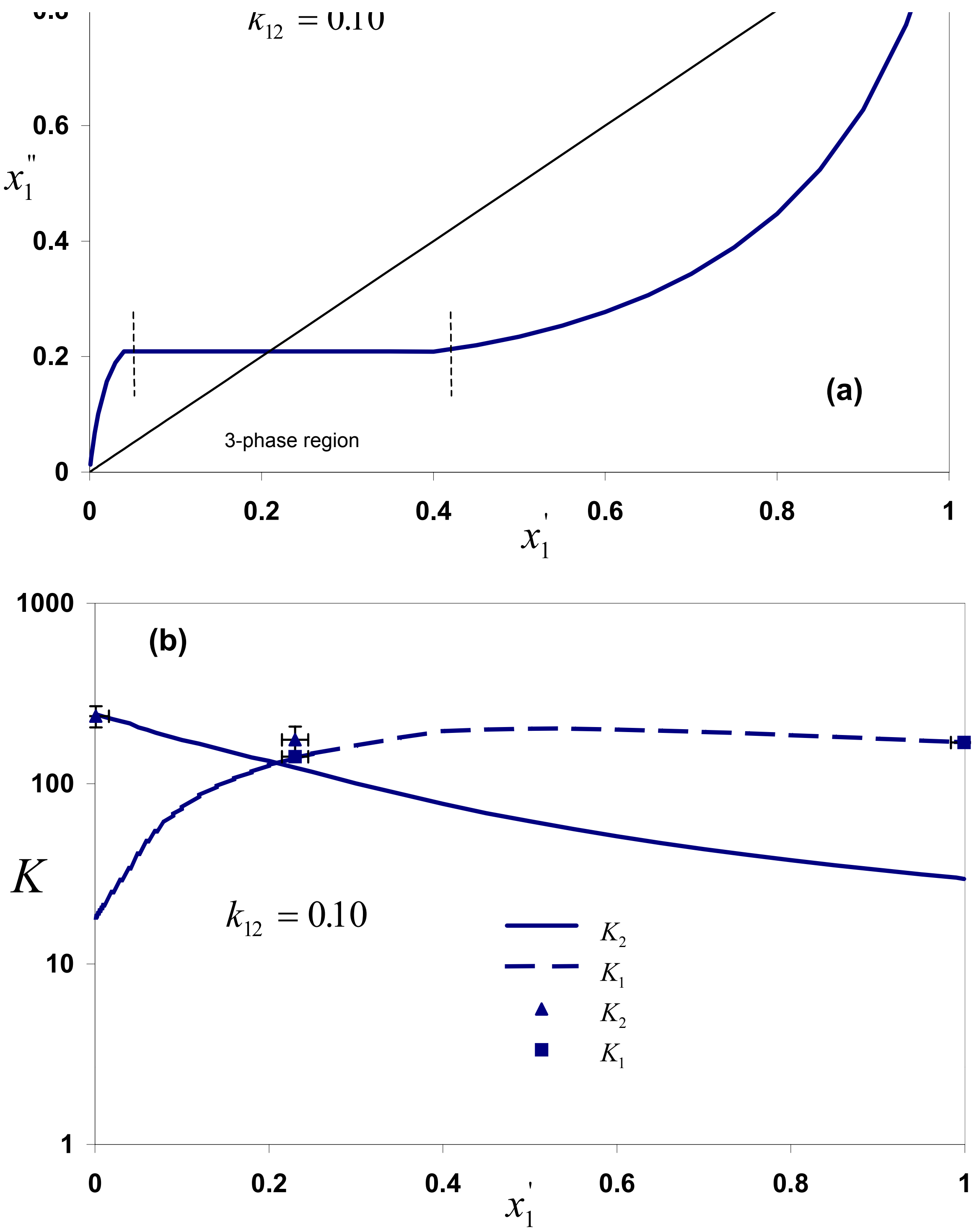

Figure 5 von Solms et al

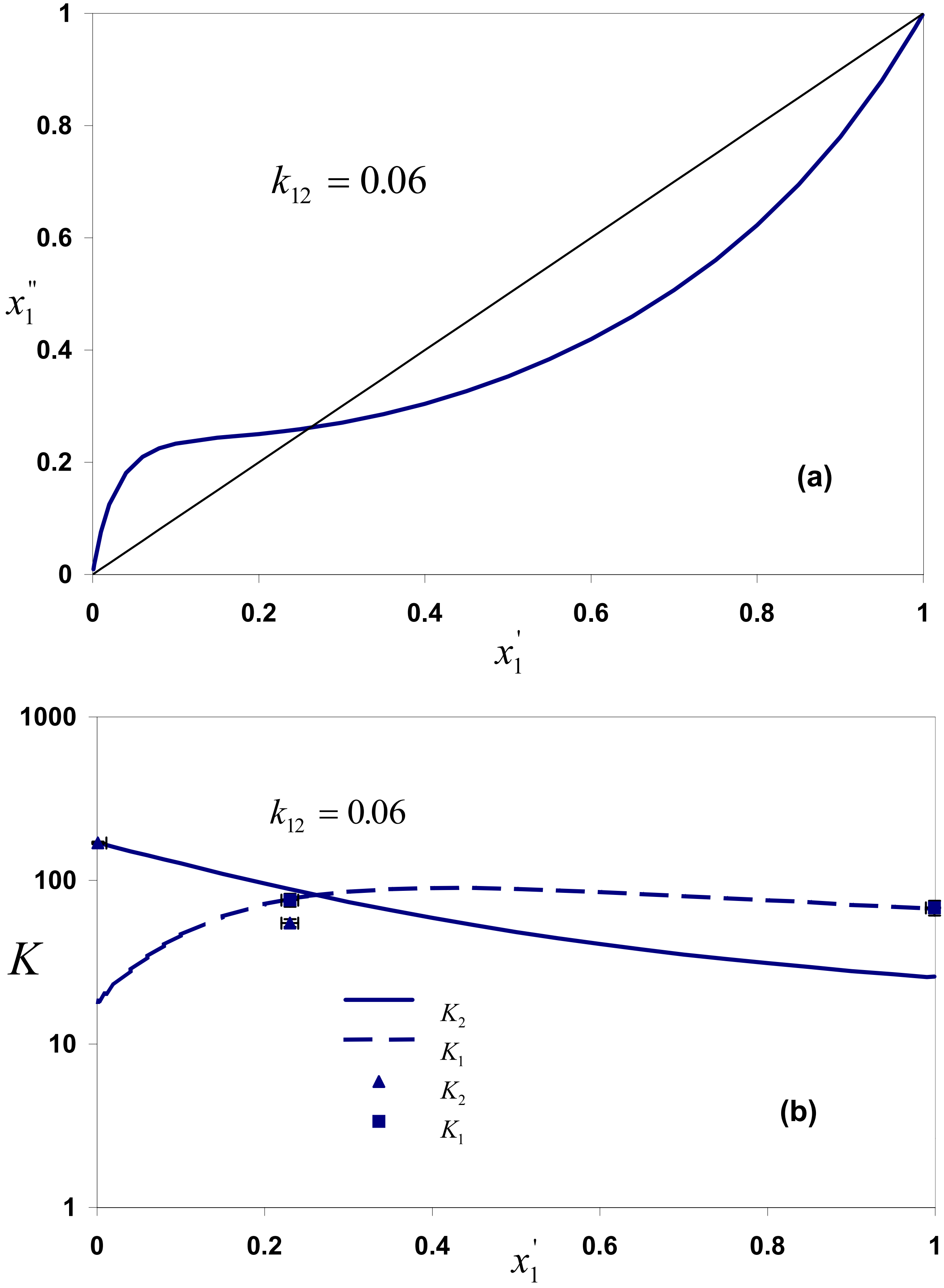

Figure 4 von Solms et al

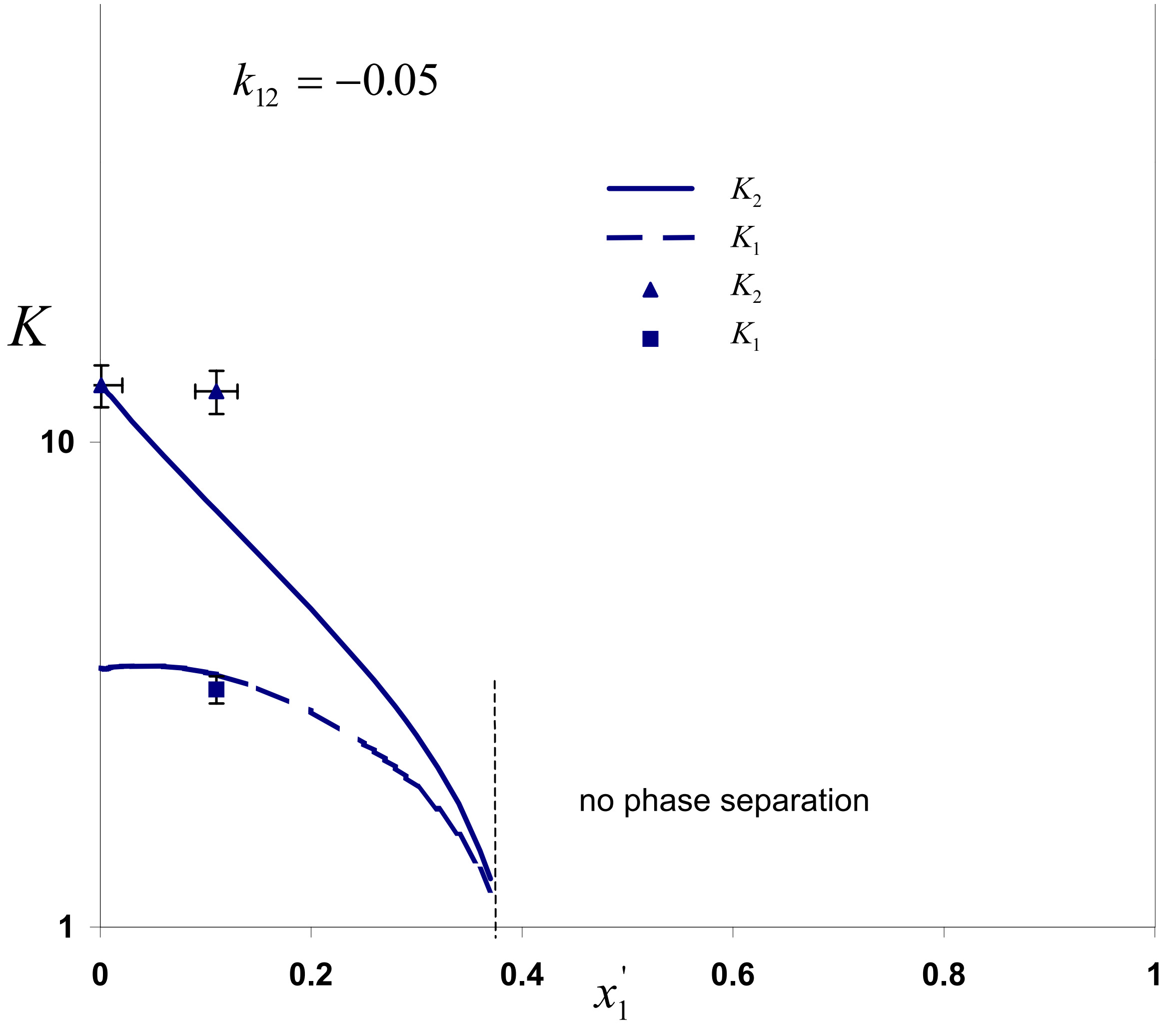

Figure 3 von Solms et al