



Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

CHEMICAL SCIENCES DIVISION

Submitted to AIChE Journal

Molecular Thermodynamics for Salt-Induced Protein Precipitation

Y.C. Chiew, D. Kuehner, H.W. Blanch, and J.M. Prausnitz

May 1994



REFERENCE COPY
Does Not Circulate
Bldg. 50 Library.
Copy 1

LBL-35587

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.

Molecular Thermodynamics for Salt-Induced Protein Precipitation

Y. C. Chiew¹, D. Kuehner, H. W. Blanch and J. M. Prausnitz

Department of Chemical Engineering

University of California

and

Chemical Sciences Division

Lawrence Berkeley Laboratory

University of California

Berkeley, CA 94720, U.S.A.

May 1994

This work was supported by the National Science Foundation under Grant BCS 9214653 and 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.

¹Department of Chemical and Biochemical Engineering
Rutgers University, Piscataway, NJ 08855-0909

ABSTRACT

A molecular-thermodynamic model is developed for salt-induced protein precipitation; the model considers the aqueous solution of globular protein molecules as a pseudo one-component system containing macroions which interact through electrostatic charge-charge repulsion, dispersive attraction, hydrophobic interactions, and forces arising from ion-excluded volume. Forces from ion-excluded volume take into account formation of ion pairs and ionic clusters at high salt concentrations; they are calculated in the context of the Percus-Yevick integral-equation theory. Hydrophobic interactions between exposed non-polar amino-acid residues on the surfaces of the protein molecules are modeled as short-range, attractive interactions between "spherical caps" on the surfaces of the protein polyions. An equation of state is derived using perturbation theory. From this equation of state we calculate liquid-liquid equilibria, i.e., equilibrium between an aqueous phase dilute in protein and another aqueous phase rich in protein; the latter represents "precipitated" protein. In the equation of state, center-to-center, spherically symmetric macroion-macroion interactions are described by the random-phase approximation, while the orientation-dependent short-range hydrophobic interaction is incorporated through the perturbation theory of associating fluids. The results obtained here suggest that either free-volume or hydrophobic-bonding effects can precipitate proteins in aqueous solutions with high salt concentrations.

1. INTRODUCTION

In research laboratories and in the biochemical industries, precipitation is commonly used to separate and isolate proteins from solutions. This technique has been applied to the recovery of proteins such as insulin, diagnostic enzymes, human growth hormone, and interferon (McGregor, 1983). Separation is achieved through addition of precipitating agents such as inorganic salts at high concentrations, nonionic polymers, polyelectrolytes, and organic solvents. (See, for example, Foster et al., 1973; Haire et al., 1984; Shih et al., 1992).

In most previous studies, protein precipitation in concentrated salt solutions has been understood as phase separation resulting in a solid phase (i.e. the protein precipitate) and a saturated protein liquid phase. Traditionally, quantitative characterization has been expressed through the protein solubility, i.e., the protein concentration in the equilibrium liquid phase. Experimental data indicate that, at fixed temperature, the concentration of protein in the liquid phase is a function of protein size and concentration, electrolyte concentration and type, pH of the solution (i.e., net charge on the protein), and interactions between exposed hydrophobic residues on the surface of the protein (Arakawa and Timasheff, 1982; 1984).

However, recent experimental results (Shih et al., 1992) on bovine serum albumin and α -chymotrypsin suggest that salt-induced protein precipitation may be more appropriately viewed as phase separation resulting in two fluid phases: a light (supernatant) fluid phase lean in protein, in equilibrium with a dense (precipitate) fluid phase rich in protein but also containing appreciable amounts of water and salt. According to this view, the degree of separation is appropriately characterized not by the protein concentration in the light phase but by the distribution coefficient, K_e , which is defined as

the ratio of the protein concentration in the dense precipitate phase to that in the light supernatant phase.

To establish a rational basis for designing a protein-precipitation process, it is useful to develop a model to provide a theoretical framework for interpretation and correlation of protein-precipitation data. The apparent solubility of a protein has been successfully correlated by the Cohn equation, which gives a simple relation between the protein concentration in the light phase and electrolyte ionic strength (Melander and Horvath, 1977; Arakawa and Timasheff, 1984). Melander and Horvath showed that the functional form of the Cohn equation may be interpreted on the basis of solvophobic effects. Recent theoretical studies have been directed at developing more fundamental models that account for the diverse interactions between the constituents in the protein solution on a molecular level. For example, Mahadevan and Hall (1990, 1992) presented a model, based on Barker-Henderson perturbation theory, for protein precipitation by a nonionic polymer. Vlachy, Blanch and Prausnitz (1993) describe a model for liquid-liquid phase separation for solutions of colloids and globular proteins, based on the random-phase approximation. However, these recent theoretical studies are concerned with aqueous solutions where the electrolyte concentration is less than 0.1 molar. Experimental studies clearly show that protein precipitation by salts requires electrolyte concentration in the range 1-10 molar.

This work presents a molecular-thermodynamic model for protein precipitation by inorganic salts. Particular attention is given to highly concentrated salt solutions. The procedure employed here represents the ternary solution (protein, electrolyte, and water) as a pseudo-one-component system containing only a continuous solvent and globular protein molecules. The solvent is an aqueous salt solution. The effect of the solvent on protein-protein interactions is taken into account through the strong influence that it exerts on the

following: electrostatic (charge-charge) repulsion, dispersive attraction, ion-excluded-volume attractive forces, and interactions between exposed hydrophobic groups on the surfaces of two or more protein molecules. Despite its simplicity, the one-component representation has been successful in explaining some experimental properties of colloidal dispersions (Grimson, 1983) and globular-protein solutions (Vlachy, et al.; 1992). A powerful advantage of this representation is that final results are based on analytical solutions to statistical-mechanical theories.

Derivations of effective potentials are discussed in Section 2. Section 3 presents a derivation of the molecular-thermodynamic equation-of-state model for protein solutions based on the random-phase approximation and the perturbation theory of association. Section 4 gives some results of model calculations.

2. PROTEIN-PROTEIN INTERACTION POTENTIALS

In the one-component model, aqueous solutions of globular proteins are represented by an assembly of spherical macroions which interact via effective solvent-dependent potentials. The potential of mean force, $W(r)$, between protein molecules (with diameter d_p and net charge z_p) is given by the sum of five potentials:

$$W(r) = W_{hs}(r) + W_{elec}(r) + W_{disp}(r) + W_{ex}(r) + W_{hy}(r, \omega) \quad (1)$$

where r is the center-to-center distance. Here, $W_{hs}(r)$ is the hard-sphere potential, $W_{elec}(r)$ is the screened Coulombic potential due to electrostatic repulsion, $W_{disp}(r)$ is the attractive dispersion potential, and $W_{ex}(r)$ is the potential due to excluded-volume effects. These four terms are spherically symmetric center-to-center potentials. $W_{hy}(r, \omega)$ represents the

interactions between exposed hydrophobic groups on the surfaces of the proteins; this potential depends not only on r but also on molecular relative orientation, indicated by ω .

Hard-sphere and electrostatic repulsions between proteins are represented by:

$$W_{hs}(r) = \infty \quad r \leq d_p \quad (2a)$$

$$= 0 \quad r > d_p \quad (2b)$$

$$\frac{W_{elec}(r)}{kT} = B \frac{\exp(-\kappa r)}{r} \quad r > d_p \quad (2c)$$

where $B = z_p^2 L_B \exp(\kappa d_p) / (1 + \kappa d_p / 2)^2$ and Bjerrum length $L_B = \beta e^2 / 4\pi\epsilon_0\epsilon_r$. In eq. (2c) $z_p e$ is the charge on the polyion, $\beta = 1/kT$, k denotes Boltzmann's constant, $\epsilon_0\epsilon_r$ represents the dielectric permittivity of the solvent, and κ^{-1} is the Debye screening length, where $\kappa^2 = 8\pi L_B N_A I$ and ionic strength $I = 0.5(z_{an}^2 \rho_{an} + z_{cat}^2 \rho_{cat})$; N_A is Avogadro's number, and z_{an} and z_{cat} are the anion and cation valences, respectively; and ρ_{an} and ρ_{cat} are the ionic number densities.

The attractive dispersion interaction $W_{disp}(r)$ is given by the following expression (Verwey and Overbeek, 1948):

$$W_{disp}(r) = -\frac{H}{12} \left\{ \frac{d_p^2}{r^2 - d_p^2} + \frac{d_p^2}{r^2} + 2 \ln \left(1 - \frac{d_p^2}{r^2} \right) \right\} \quad \text{for } r > d_p, \quad (3)$$

where H represents the Hamaker constant. For large values of r , eq. (3) reduces to

$$W_{disp}(r) = -\frac{H}{36} \left(\frac{d_p}{r} \right)^6 \quad \text{for } r \gg d_p. \quad (4)$$

Eq. (4) is the simplified large- r limit for $W_{disp}(r)$. This form of the dispersion potential has been used successfully in modeling phase separation of colloidal systems (Grimson, 1983). We recognize that, as r approaches d_p , eq. (4) underestimates the contribution of dispersion forces to the total potential of mean force. However, upon using eq. (4) instead of eq. (3), we can obtain an analytic equation of state for calculating phase equilibria.

Ion-Excluded-Volume Potential

The term $W_{ex}(r)$ in eq. (1) accounts for the interaction between a pair of protein macroions due to the excluded volume of ions in solution. The literature has reported experimental and theoretical studies of the effect of volume exclusion by solvent or small solutes on macromolecular interactions (Israelachvili, 1985; Henderson and Lozada-Cassou, 1986; Henderson, 1988). Henderson and coworkers computed the potentials of mean force for two large rigid hard spheres immersed in a one-component hard-sphere fluid (Henderson and Lozada-Cassou, 1986; Henderson, 1988) and the adhesive hard-sphere fluid (Jamnik et al, 1991), in the context of the Percus-Yevick (PY) integral-equation theory. Their results reveal that the PY theory provides semi-quantitative description of experimental results (Henderson, 1992).

For concentrated salt systems typically used in protein precipitation, ion pairs and larger ionic clusters are expected to form in the electrolyte solution (Robinson and Stokes, 1959). These ionic aggregates have an important effect on the ion-excluded-volume potential of mean force, $W_{\alpha}(r)$, between protein macroions because the diameter of an ion cluster is appreciably larger than that of a single ion; cluster formation increases the ion-excluded-volume effect. This potential of mean force can be obtained from the radial distribution function $g_{33}(r)$ between two protein particles (denoted by subscript 3),

separated by a center-to-center distance r , in a system of ionic clusters. In our model, $W_{ex}(r)$ is obtained by solving for $g_{33}(r)$, at infinite dilution, for two large hard spheres immersed in a binary system of much smaller ions (denoted by subscripts 1 and 2), with diameters d_1 and d_2 , and number densities ρ_1 and ρ_2 . The interactions between like ionic species, i.e., 1-1 and 2-2 interactions, follow the simple repulsive hard-sphere potential. Cation-anion interactions are described by the adhesive hard-sphere potential, which leads to formation of bonds between unlike species; the bonded species are ionic clusters. The adhesive hard-sphere (AHS) model (Baxter, 1968b) is defined below in terms of the total and direct correlation functions, $h_{ij}(r)$ and $c_{ij}(r)$, respectively. This simple adhesive hard-sphere system is able to model aggregation of small "ionic" particles; it is used to mimic the formation of ion pairs and ionic clusters from anions and cations in the electrolyte solution. This simple calculation assumes that the electrostatics between ions are adequately taken into account through the electrostatic potential $W_{elec}(r)$. The potential of mean force $W_{ex}(r)$ can be calculated from the protein-protein radial distribution function $g_{33}(r)$ from the following equation (McQuarrie, 1975):

$$W_{ex}(r; \rho_1, \rho_2) = -kT \ln[g_{33}(r; \rho_1, \rho_2)] = -kT \ln[h_{33}(r; \rho_1, \rho_2) + 1] \quad (5)$$

where $h_{33}(r)$ is the total correlation function in the limit $\rho_3 = 0$.

To obtain $W_{ex}(r)$, we consider the three-component system which consists of species 1 (ion), 2 (ion) and 3 (protein). The function $g_{33}(r)$ is obtained by solving the set of multi-component Ornstein-Zernike (OZ) integral equations

$$h_{ij}(r) = c_{ij}(r) + \sum_k \rho_k \int c_{ik}(s) h_{kj}(r-s) ds \quad (6)$$

for finite values of ρ_1 and ρ_2 and in the limit $\rho_3=0$. To obtain a solution, we use the Percus-Yevick (PY) approximation. For the adhesive hard-sphere system considered here, the PY theory provides the following boundary conditions to the OZ equation:

$$h_{ij}(r) = -1 + \frac{\lambda_{ij}d_{ij}}{12} \delta(r - d_{ij}) \quad 0 < r < d_{ij} \quad (7)$$

and

$$c_{ij}(r) = 0 \quad r > d_{ij} \text{ (for all } ij \text{ pairs)} \quad (8)$$

where $\lambda_{ij} = 0$ for $ij \neq 12$ or 21 . Due to symmetry, $\lambda_{12} = \lambda_{21}$. Equation (7) implies that the interactions between species 1-1, 2-2, 3-3, 1-3 and 2-3 are determined by hard-sphere potentials. Clustering between species 1 and 2 is explicitly accounted for by the Dirac delta function $\delta(r - d_{12})$ which is defined so that $\int_0^{\infty} f(r)\delta(r - d_{12})dr = f(r = d_{12})$. Parameter λ_{12} takes into account the average number of 1-2 direct contacts that are formed in the system; it is related to N_{12} , the average number of species-2 ions that have direct bonds with a species-1 ion by

$$N_{12} = 4\pi\rho_2 \int_0^{d_{12}} r^2 g_{12}(r) dr = 2\eta_2 \lambda_{12} \left(\frac{d_{12}}{d_2}\right)^3 \quad (9)$$

where $\eta_2 = \pi\rho_2 d_2^3 / 6$. When $N_{12} = 1$, each species-1 ion, on the average, has one species-2 bonded neighbor. When the electrolyte is a symmetric salt, $N_{21} = N_{12}$. However, for asymmetric electrolytes, $N_{21} = N_{12}(z_2/z_1)$, where z_1 and z_2 are the ion valences.

Because of the short-range nature of $c_{ij}(r)$, the OZ equation may be solved through the use of the Wiener-Hopf factorization technique (Baxter, 1968a). The resulting equation for the protein-protein total correlation function $h_{33}(r)$ is decoupled from the total

correlation function $c_{33}(r)$, and can be shown to follow the equation below (Perram and Smith, 1977; Barboy and Tenne, 1979; Chiew, 1991):

$$rh_{33}(r) = -q'_{33}(r) + 2\pi \sum_{k=1}^2 \rho_k \int_{S_k}^{d_k} dt q_{3k}(t) (r-t) h_{k3}(r-t) \quad (10)$$

where $S_{ij} = (d_{ii}-d_{jj})/2$. The functions $h_{13}(r)$ and $h_{23}(r)$ represent the protein-ion total correlation functions; they are obtained from

$$rh_{i3}(r) = -q'_{i3}(r) + 2\pi \sum_{k=1}^2 \rho_k \int_{S_k}^{d_k} dt q_{ik}(t) (r-t) h_{k3}(r-t). \quad (11)$$

In the above equations, function $q_{ij}(r)$ is given by

$$q_{ij}(r) = \frac{a_i}{2}(r^2 - d_{ij}^2) + b_i(r - d_{ij}) + \frac{\lambda_{ij}d_{ij}^2}{12} \quad \text{for } S_{ij} < r < d_{ij} \quad (12a)$$

$$= 0 \quad \text{for } r > d_{ij} \quad (12b)$$

and

$$q'_{ij}(r) = a_i r + b_i \quad \text{for } S_{ij} < r < d_{ij} \quad (13a)$$

$$= 0 \quad \text{for } r > d_{ij} \quad (13b)$$

where $\lambda_{ij} = 0$ for $ij \neq 12$ or 21 . The PY solution provides the following expressions for parameters a_i and b_i (for $i = 1, 2$ and 3):

$$a_i = \frac{1-H+3d_{ii}G}{(1-H)^2} - \frac{Y_1}{(1-H)} \delta_{ii} - \frac{Y_2}{(1-H)} \delta_{i2} \quad (14a)$$

$$b_i = -\frac{3d_{ii}^2G}{2(1-H)^2} + \frac{d_{ii}Y_1}{2(1-H)} \delta_{ii} + \frac{d_{ii}Y_2}{2(1-H)} \delta_{i2} \quad (14b)$$

$$Y_1 = \eta_2 \lambda_{12} \left(\frac{d_{12}}{d_2} \right)^2 \quad (14c)$$

$$Y_2 = \eta_1 \lambda_{12} \left(\frac{d_{12}}{d_1} \right)^2 \quad (14d)$$

$$G = \frac{\pi}{6} \sum_{k=1}^2 \rho_k d_{kk}^2 \quad (14e)$$

$$H = \frac{\pi}{6} \sum_{k=1}^2 \rho_k d_{kk}^3 \quad (14f)$$

and the inequality

$$\frac{1+2H}{(1-H)^2} + \frac{\eta_1 \eta_2 \lambda_{12} (6 - \lambda_{12})}{(1-H)^3} > 0. \quad (14g)$$

Here δ_{ij} is the Kronecker delta, i.e., $\delta_{ij} = 1$ for $i = j$, and zero otherwise. The second and third terms on the right hand sides of eqs. (14a) and (14b) vanish if $i \neq 1$ and 2, respectively. Eqs. (12) through (14a-14f) are the PY analytic solution of the OZ integral equation, i.e., eq. (6), subject to the boundary conditions given by eqs. (7) and (8). The PY solution further requires that the inequality given by eq. (14g) must be satisfied to ensure physically admissible solutions (Baxter, 1968; Barboy and Tenne, 1979); this means that λ_{12} or N_{12} depend on the density of the system and must be properly chosen for model calculations. The total correlation function $h_{33}(r)$ is obtained by first calculating $h_{13}(r)$ and $h_{23}(r)$ from eq.(11), followed by solving $h_{33}(r)$ from eq.(10); these calculations can be performed using the simple numerical procedure proposed by Perram (1975). That calculation is much simpler than solving the set of multi-component Ornstein-Zernike integral equations (i.e., eq. (6)) simultaneously through a Fourier-transform technique. The potential $W_{ex}(r)$ is related to $h_{33}(r)$ by eq. (5); it is independent of protein concentration, and depends only on the number densities of ions (species 1 and 2), ion diameters d_1, d_2 , and parameter N_{12} that characterizes the degree of ion clustering. Densities ρ_1 and ρ_2 are related by the stoichiometric relation $z_1 \rho_1 + z_2 \rho_2 = 0$, where z_1 and z_2 are the valences of the ions.

Figure 1 shows the computed potential of mean force $W_{ex}(r)/kT$ plotted as a function of protein-protein center-to-center distance r for a salt solution containing ions with diameters $d_1 = 4.6\text{\AA}$ and $d_2 = 2.96\text{\AA}$, $|z_1| = 2$, $|z_2| = 1$, at a salt concentration $C_s = 0.5$ moles/liter, with $d_p = 40\text{\AA}$, for different values of N_{12} . Potential $W_{ex}(r)/kT$ is attractive and the range of the potential increases with N_{12} , the degree of ionic clustering. This behavior implies that at high salt concentrations, when ionic pairs and clusters are formed, the interaction between protein macroions becomes increasingly attractive; it is this attraction that has a major influence on phase separation.

In this approximate calculation for the effect of the excluded-volume interaction of ions on $g_{33}(r)$, we have assumed that the electrostatics in the system are adequately taken into account by the screened Coulombic potential $W_{elec}(r)$ given in eq. (2c). Thus, at high salt concentrations, since the Debye parameter κd_p is large, electrostatic interactions are essentially eliminated and should play a minor role in phase separation. In practice, however, as proteins approach contact, the discrete nature of the charges on the surface of the protein molecules may give rise to electrostatic interactions that cannot be accurately described by a simple screened Coulombic repulsion term. Since it has been observed experimentally that high salt concentrations are required to bring about phase separation, the contribution of anion and cation clustering to the excluded-volume interaction $W_{ex}(r)$ plays an important role in the phase separation of proteins; this role is explicitly taken into account through the adhesive hard-sphere model potentials outlined above.

Hydrophobic Interactions

The hydrophobic interaction between exposed non-polar amino acid residues on the surfaces of the protein molecules is, in general, attractive, short-range, and orientation-dependent. Hydrophobic bonds are formed when two hydrophobic groups come into contact with each other, and cause association or aggregation of protein molecules in the system. In this work, hydrophobic interaction is represented by a potential model used for associating fluids (Jackson et al, 1988). The shapes of these hydrophobic groups are idealized as "circular patches" or "spherical caps" located on the surface of the macroion. As indicated in Figure 2, interaction potential $W_{hy}(r_{12})$ between hydrophobic "patch" A on the surface of particle 1 and hydrophobic "patch" B on the surface of particle 2 is defined such that

$$W_{hy}(r_{12}) = -\varepsilon \quad ; \text{ if } |r_{12}| \leq r_{12,c} \text{ and } d_{1,A} \cdot r_{12} \leq \cos\theta_{1,c} \text{ and } -d_{2,B} \cdot r_{12} \leq \cos\theta_{2,c}$$

$$= 0 \quad ; \text{ otherwise} \quad (15)$$

where d_{iK} represents the vector joining the center of particle i to the center of patch K (located on the surface of particle i). The quantity r_{12} denotes the vector joining the center of molecule 1 to the center of molecule 2. The vector dot product $d_{1,A} \cdot r_{12} = \cos\theta_1$ where angle θ_1 denotes the angle between r_{12} and $d_{1,A}$. Angle θ_2 is defined in a similar manner. Hence, two hydrophobic "patches" on two different particles are considered to form a "bond" if the centers of the two particles are within a distance $r_{12,c}$ from each other, and the two hydrophobic groups satisfy the orientation constraints $d_{1,A} \cdot r_{12} \leq \cos\theta_{1,c}$ and $-d_{2,B} \cdot r_{12} \leq \cos\theta_{2,c}$. The quantities $r_{12,c}$, $\theta_{1,c}$, and $\theta_{2,c}$ characterize the range of the hydrophobic interaction and the sizes of the hydrophobic groups.

3. EQUATION OF STATE

Having established and defined pertinent potentials of mean force, it is now necessary to construct a molecular-thermodynamic model which relates these potentials to macroscopic thermodynamic properties. In this work, that model is based on perturbation theory. The center-to-center spherically symmetrical electrostatic, dispersion and excluded-volume interactions are incorporated into the model in the context of the random-phase approximation. The orientation-dependent hydrophobic interaction is included through the first-order perturbation theory of associating fluids formulated by Wertheim (1986, 1987).

The Random Phase Approximation (RPA) has been used previously to model the phase transition and structure factor of colloids (Grimson, 1983), and to describe liquid-liquid phase separation of proteins due to addition of polymers (Mahadevan and Hall, 1990, 1992; Vlachy et al, 1993). In the RPA, an assembly of hard spheres is used as the reference system, while the remaining spherically symmetric interactions are treated as perturbations. Following Grimson (1983) and Vlachy, et. al (1993), the compressibility factor Z_{sym} due to spherically symmetric potentials within the RPA is expressed as

$$Z_{sym} = \frac{P}{\rho kT} = Z_{hs} + Z_{pert} = \frac{P_{hs}}{\rho kT} + \frac{\rho U}{2kT} \quad (16)$$

Here P_{hs} represents the pressure of the hard-sphere fluid; $\rho = \rho_p$ is the number density of protein molecules; and the energy per unit density U is given by

$$U = 4\pi \int (W_{elec}(r) + W_{disp}(r) + W_{ex}(r)) r^2 dr. \quad (17)$$

Because U is an energy per unit density, it is assumed to be independent of protein density; U depends only on potentials of mean force between protein molecules. The residual

Helmholtz energy per molecule a^{res} / kT is defined as $(a/kT)_{total} - (a/kT)_{ideal\ gas}$ at the same temperature and density; it is given by

$$\frac{a^{res}}{kT} = \frac{a_{hs}^{res}}{kT} + \frac{a_{pert}^{res}}{kT} = \frac{a_{hs}^{res}}{kT} + \frac{\rho U}{2kT}. \quad (18)$$

For the pressure and residual Helmholtz energy of the hard-sphere reference system, the Carnahan-Starling expressions are used,

$$\frac{P_{hs}}{\rho kT} = \frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3} \quad (19)$$

and

$$\frac{a_{hs}^{res}}{kT} = \frac{4\eta - 3\eta^2}{(1 - \eta)^2} \quad (20)$$

where the volume fraction of protein is $\eta = \eta_p \equiv \pi \rho_p d_p^3 / 6$. Energy per unit density U is found from eq. (17); it is given by

$$\frac{\rho U}{kT} = 8\eta z_p^2 L_B \frac{[1 + 3/(\kappa d_p) + 3/(\kappa d_p)^2]}{d_p(1 + \kappa d_p/2)^2} - \frac{4\eta H}{9kT} + \frac{\rho U_{ex}}{kT} \quad (21)$$

where $U_{ex} = 4\pi \int W_{ex}(r)r^2 dr$ is the contribution to U from the ion-excluded-volume interaction.

The contribution of the non-spherically-symmetric hydrophobic interaction to the residual Helmholtz energy and pressure of the system are evaluated using the first-order perturbation theory of associating fluids formulated by Wertheim (1986, 1987), extended to mixtures by Gubbins and co-workers (Jackson et al, 1988). At a given protein or particle number density ρ and temperature T , this theory gives the Helmholtz energy of the

associating system relative to that of the non-associating reference system. The reference system is an assembly of non-aggregating protein macroions that interact through the spherically-symmetric potentials. For a protein molecule consisting of M equivalent or identical exposed hydrophobic sites interacting via the potential given by eq. (15), the first-order perturbation theory yields the following contribution to the Helmholtz energy due to hydrophobic association (Jackson et al, 1988):

$$\frac{a^{assoc}}{kT} = M \left[\ln X - \frac{X}{2} + \frac{1}{2} \right]. \quad (22)$$

Here,

$$X = \frac{-1 + \sqrt{1 + 4M\rho\Delta}}{2M\rho\Delta} \quad (23a)$$

$$\Delta = g_{pp}(d_p) [\exp(\varepsilon / kT) - 1] V \quad (23b)$$

$$V = 3 \frac{[1 - \cos(\theta_{1c})][1 - \cos(\theta_{2c})] (r_{12,c} - d_p)}{2\pi d_p} \quad (23c)$$

where M is the number of hydrophobic groups on the protein surface, MX is the fraction of hydrophobic sites that are not bonded, $g_{pp}(d_p)$ is the radial distribution function of the non-aggregating system, ε / kT represents the characteristic energy of the hydrophobic interaction, and the dimensionless parameter V corresponds to the volume of interaction between two attractive square-well hydrophobic sites divided by the volume of a single protein particle. Because we do not know $\theta_{1,c}$ and $\theta_{2,c}$, which characterize the size of exposed hydrophobic groups, and $r_{12,c}$, which characterizes the range of hydrophobic interaction, parameter V , given by eq. (23c), can be used to quantify the "hydrophobic association" volume. Hence, the hydrophobic interaction is characterized by two model

parameters, viz. association energy ϵ and association volume V . The function $g_{pp}(d_p)$ for the non-agglomerating protein macroions (which interact through the hard-sphere, electrostatic, dispersion and ion-excluded-volume potentials) is estimated using the EXP approximation (Konior and Jedrzejek, 1985):

$$g_{pp}(d_p) = g_{hs}(d_p) \cdot \exp[-(W_{elec}(d_p) + W_{disp}(d_p) + W_{ex}(d_p))/kT] \quad (24)$$

where $g_{hs}(d_p)$ is given by the contact value of the Carnahan-Starling expression for the hard-sphere radial distribution function:

$$g_{hs}(d_p) = \frac{1}{(1-\eta)} + \frac{3}{2} \frac{\eta}{(1-\eta)^2} + \frac{1}{2} \frac{\eta^2}{(1-\eta)^3}. \quad (25)$$

The first-order perturbation theory of association assumes that the interactions between hydrophobic sites on different molecules are independent of each other and that no ring structures (only tree-like structures) are formed in the protein aggregates. Combining contributions to the residual Helmholtz energy and compressibility factor from spherically symmetric and non-symmetric interactions, i.e., eqs. (16), (18), (19), (20) and (22), it follows that a^{res}/kT and Z are given by

$$\frac{a^{res}}{kT} = \frac{4\eta - 3\eta^2}{(1-\eta)^2} + \frac{\rho U}{2kT} + M \left[\ln X - \frac{X}{2} + \frac{1}{2} \right] \quad (26)$$

and

$$Z = \frac{P}{\rho kT} = \frac{1 + \eta + \eta^2 - \eta^3}{(1-\eta)^3} + \frac{\rho U}{2kT} + M \left\{ \frac{1}{X} - \frac{1}{2} \right\} \eta \left(\frac{\partial X}{\partial \eta} \right)_T \quad (27)$$

The chemical potential μ/kT can be obtained from eqs. (26) and (27) through the thermodynamic relation $\mu^{res}/kT = a^{res}/kT + Z - 1$, and the ideal-gas chemical potential, $\mu^{ig}/kT = \mu'/kT + \ln \rho$, where μ' is a function only of temperature. We obtain

$$\frac{\mu - \mu'}{kT} = \ln \rho + \frac{\mu_{hs}}{kT} + \frac{\rho U}{kT} + M \left\{ \ln X - \frac{X}{2} + \frac{1}{2} + \eta \left(\frac{\partial X}{\partial \eta} \right)_T \left[\frac{1}{X} - \frac{1}{2} \right] \right\}. \quad (28)$$

Here,

$$\frac{\mu_{hs}}{kT} = \frac{\eta(8 - 9\eta + 3\eta^2)}{(1 - \eta)^3} \quad (29)$$

and $\mu'/kT = \ln(\Lambda^3)$ where the de Broglie wavelength $\Lambda = h/(2\pi mkT)^{1/2}$. Energy per unit density U , given by eq. (21), gives the effects of electrostatic, dispersion, and ion-excluded-volume interactions. At equilibrium, protein concentrations in the supernatant and dense-fluid phases are calculated from eqs. (27) and (28) based on the classical equilibrium conditions:

$$\mu_s = \mu_d \quad (30a)$$

$$P_s = P_d \quad (30b)$$

Here, subscripts s and d denote the equilibrium supernatant and dense protein phases, respectively.

4. RESULTS AND DISCUSSION

We first examine the effect of salt concentration on the phase behavior of the system. Figure 3 shows the reduced pressure Pv_0/kT , computed from eq. (27), plotted as a function of the protein volume fraction η for $H/kT = 9.6$, $|z_{an}| = 2$, $|z_{cat}| = 1$, $|z_p| = 5$, $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, $N_{12} = 1$, and different values of salt concentrations C_s , in the absence of hydrophobic interactions. Here v_0 represents the volume of a single protein molecule. The pressure increases monotonically with protein volume fraction η at the two low salt concentrations. However, when $C_s = 1.6$ moles/liter, the pressure curve exhibits a van der Waals loop, indicating that the system undergoes a fluid-fluid phase transition. This result suggests that rising electrolyte concentrations increase the excluded volume attraction between proteins, leading to phase separation. Further, distribution coefficient K_e increases with N_{12} , a measure of the degree of ionic clustering. Because ionic clustering enhances attraction in $W_{ex}(r)$, clustering increases the protein concentration in the precipitate phase and lowers the protein concentration in the supernatant phase.

The distribution coefficient K_e of the protein system can be obtained from the equilibrium conditions, i.e., eqs. (30a) and (30b), through the use of eqs. (27) and (28). Distribution coefficient K_e is given by the ratio of the equilibrium number density of protein in the dense phase to that in the supernatant phase. Neglecting hydrophobic interactions, Figure 4 shows predicted distribution coefficients K_e plotted as a function of ionic strength for $H/kT = 9.6$, $|z_{an}| = 2$, $|z_{cat}| = 1$, $|z_p| = 5$, $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, and two values of N_{12} . Coefficient K_e increases monotonically with electrolyte concentration.

Again neglecting hydrophobic interactions, Figure 5 shows the variation of distribution coefficient K_e as a function of z_p , the net charge of protein, for $H/kT = 8$, $|z_{an}| = 2$, $|z_{cat}| = 1$, $|z_p| = 5$, $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, $N_{12} = 1$, and $I = 6.0$ and 6.6

moles/liter. Distribution coefficient K_e is insensitive to z_p because charge-charge electrostatic repulsion between protein molecules is screened out in highly concentrated salt solutions. Since the protein charge z_p is directly related to the pH of the solution, this prediction further implies that pH has little influence on phase separation in our model. In contrast, Figure 5 shows that ionic strength has a strong impact on protein precipitation. At a fixed value of z_p , distribution coefficient K_e increases by nearly a factor of two as the ionic strength increases from 6.0 to 6.6 moles/liter. This increase in K_e is due to the enhanced effect of the ion-excluded-volume contribution.

We now present calculations on the effect of hydrophobic interactions on protein phase separation. The contribution of the hydrophobic interaction to the thermodynamic properties of the system is primarily characterized by M , the number of hydrophobic sites, and ϵ/kT and V , the characteristic energy and volume of hydrophobic attraction, respectively. Figure 6 shows distribution coefficients K_e as a function of ϵ/kT , for $H/kT = 6$, $N_{I2} = 1$, $|z_{cat}| = 1$, $|z_{an}| = 1$, $|z_p| = 8$, $d_{an} = d_{cat} = 3.4\text{\AA}$, $d_p = 40\text{\AA}$, $V = 0.006$, and $I = 2.5$ moles/liter, for two different values of M . Distribution coefficient K_e increases with ϵ/kT , because as the strength of hydrophobic attraction rises, protein molecules tend to form aggregates. Protein aggregation raises the protein concentration in the precipitate phase, and lowers the protein concentration in the supernatant phase. At a fixed ϵ/kT , distribution coefficient K_e increases with M , the average number of exposed hydrophobic sites.

We now examine the effect of Hamaker constant H/kT on the phase separation of proteins. Figure 7 shows distribution coefficient K_e as a function of ionic strength for $|z_{cat}| = 1$, $|z_{an}| = 1$, $|z_p| = 8$, $d_{an} = d_{cat} = 3.4\text{\AA}$, $d_p = 40\text{\AA}$, $V = 0.006$, $M = 4$, $\epsilon/kT = 4$, $N_{I2} = 1$, and three reduced Hamaker constants H/kT . Distribution coefficient K_e increases with rising H/kT at fixed I . Figure 8 shows distribution coefficient K_e as a function of protein

diameter d_p for $|z_{cat}| = 1$, $|z_{an}| = 1$, $|z_p| = 8$, $d_{an} = d_{cat} = 3.4\text{\AA}$, $H/kT = 6$, $V = 0.006$, $M = 4$, $\epsilon/kT = 4$, $I = 2.5$ moles/liter, and $N_{12} = 1$. Distribution coefficient increases with increasing d_p , consistent with experimental observations.

Finally, we study the variation of K_e with ion diameters. Figure 9 shows the distribution coefficient K_e as a function of ion diameter for a monovalent electrolyte with $d_{ion} = d_{an} = d_{cat}$, $d_p = 40\text{\AA}$, $|z_{cat}| = 1$, $|z_{an}| = 1$, $|z_p| = 8$, $H/kT = 4$, $M = 4$, $\epsilon/kT = 4$, $V = 0.006$, $N_{12} = 1$, and two values of I . As expected, K_e increases with the ion diameter, especially at high salt concentration, indicating once more the importance of the ion-excluded-volume contribution.

In summary, we have derived an approximate statistical-mechanical equation-of-state model for salt-induced protein precipitation. In this model, proteins are considered to be macroions which interact with electrostatic repulsion, dispersion attraction, ion-excluded-volume attraction, and hydrophobic interactions. Thermodynamic properties of the system are derived using perturbation theory. The model indicates that (i) distribution coefficient K_e is insensitive to net charge of protein due to strong electrostatic screening at high salt concentrations, (ii) concentration of the electrolyte plays a major role in affecting phase splitting in protein solutions; distribution coefficient K_e increases monotonically with salt concentration, (iii) distribution coefficient K_e is particularly sensitive and rises with ion diameters, protein diameter and ionic clustering, and (iv) aggregation of protein due to hydrophobic interaction may play an important role in the precipitation of proteins; the extent of aggregation is a strong function of M , the average number of exposed associating sites.

Precipitation of proteins by salts may result from ion-excluded-volume effects or from hydrophobic-bond aggregation, or both. Using physically reasonable parameters,

either excluded volume or aggregation can be used to interpret protein-precipitation data. The relative importance of these effects can be estimated only from experimental studies.

Acknowledgments

This work was supported by the National Science Foundation, Grant # BCS 9214653, and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the US Department of Energy. Y. C. Chiew thanks John Prausnitz for his kind hospitality when YCC spent his sabbatical leave at the University of California, Berkeley in Fall 1992.

Literature Cited

- Arakawa, T., and S. N. Timasheff, "Preferential Interactions of Proteins with Salts in Concentrated Solutions," *Biochemistry* **21**, 6545 (1982).
- Arakawa, T., and S. N. Timasheff, "Mechanism of Protein Salting-In and Salting-Out by Divalent Cation Salts: Balance Between Hydration and Salt Binding," *Biochemistry* **23**, 5912 (1984).
- Baxter, R. J., "Ornstein-Zernike Relation for a Disordered Fluid", *Aust. J. Phys.* **21**, 563 (1968).
- Baxter, R. J., "Percus-Yevick Equation for Hard Spheres with Surface Adhesion", *J. Chem. Phys.*, **49** 2770 (1968).
- Barbooy, B., and R. Tenne, "Distribution Functions and Equations of State of Sticky Hard Sphere Fluids in the Percus-Yevick Approximation," *Chem. Physics* **38**, 369 (1979).
- Chiew, Y. C., "Sol-Gel Transition in Dense Dispersions: Effect of Particle Bridging," *J. Coll. Inter. Sci.* **143**, 397 (1991).
- Carnahan, N., and K. Starling, "Thermodynamic Properties of a Rigid-Sphere Fluid," *J. Chem. Phys.* **53**, 600 (1970).
- 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).
- Grimson, M. J., "Small Angle Scattering from Colloidal Dispersions," *J. Chem. Soc. Faraday Trans. II* **79**, 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," *Biopolymers* **23**, 2761 (1984).
- Henderson, D., "An Explicit Expression for the Solvent Contribution to the Force Between Colloidal Particles Using a Hard Sphere Model," *J. Coll. Inter. Sci.* **121**, 486 (1988).
- Henderson, D., and M. Lozada-Cassou, "A Simple Theory for the Force Between Spheres Immersed in a Fluid," *J. Coll. Inter. Sci.* **114**, 180 (1986).
- Henderson, D., "The Interaction Between Macrospheres in Solution," *Fluid Phase Equil.* **76**, 1 (1992).
- Israelachvili, J. N., *Intermolecular and Surfaces Forces*, Academic Press, London, New York, 1985.
- Jackson, G., W. G. Chapman, and K. E. Gubbins, "Phase Equilibria of Associating Fluids: Spherical Molecules with Multiple Bonding Sites," *Molec. Phys.* **65**, 1 (1988).
- Jamnik, A., D. Bratko, and D. J. Henderson, "Lyophobic Interaction in Baxter's Adhesive Fluid," *J. Chem. Phys.* **94**, 8210 (1991).

Konior, J., and C. Jedrzejek, "Analytic Formulation of the WCA Type Perturbation Theory for a Hard-Core with One-Yukawa Tail Fluid," *Molec. Phys.* **55**, 187 (1985).

Mahadevan, H., and C. K. Hall, "Statistical-Mechanical Model of Protein Precipitation by Nonionic Polymer," *AIChE J.* **36**, 1517 (1990).

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

McGregor, W. C., "Large Scale Isolation and Purification of Proteins from Recombinant *E. Coli*," *Ann. N. Y. Acad. Sci.* **413**, 231 (1983).

McQuarrie, D. A., *Statistical Mechanics*, Harper-Collins, New York, 1976.

Melander, W., and C. Horvath, "Salt Effect on Hydrophobic Interactions in Precipitation and Chromatography of Proteins: An Interpretation of the Lyotropic Series," *Arch. Biochem. Biophys.* **183**, 200 (1977).

Perram, J. W., "Hard Sphere Correlation Functions in the Percus-Yevick Approximation," *Molec. Phys.* **30**, 1505 (1975).

Perram, J. W., and E. R. Smith, "Competitive Adsorption via the Percus-Yevick Approximation," *Proc. Roy. Soc. Lond.* **A353**, 193 (1977).

Robinson, R. A., and R. H. Stokes, *Electrolyte Solutions*, Academic Press, New York, 1959.

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

Verwey, E. J. W., and J. Th. G. Overbeek, *Theory of the Stability of Lyophobic Colloids*, Elsevier, 1948.

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

Wertheim, M. S., "Fluids with Highly Directional Attractive Forces. III. Multiple Attraction Sites," *J. Stat. Phys.* **42**, 459 (1986).

Wertheim, M. S., "Fluids with Highly Directional Attractive Forces. IV. Equilibrium Polymerization," *J. Stat. Phys.* **42**, 477 (1986).

Figure Captions

- Figure 1. Ion-excluded-volume potential of mean force $W_{ex}(r)/kT$ as a function of protein center-to-center distance for three values of clustering parameter N_{I2} . Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, $|z_{an}| = 2$, $|z_{cat}| = 1$; and salt concentration $C_s = 0.5$ moles/liter.
- Figure 2. Schematic of the short-range orientation-dependent hydrophobic interaction.
- Figure 3. Reduced pressure as a function of protein packing fraction η for three salt concentrations. Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, $|z_p| = 5$, $|z_{an}| = 2$, $|z_{cat}| = 1$, $H/kT = 9.6$, and $N_{I2} = 1$.
- Figure 4. Distribution coefficient K_e as a function of ionic strength I for two values of N_{I2} . Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, $|z_p| = 5$, $|z_{an}| = 2$, $|z_{cat}| = 1$, and $H/kT = 9.6$.
- Figure 5. Distribution coefficient K_e as a function of net charge of protein z_p for two ionic strengths. Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = 4.6\text{\AA}$, $d_{cat} = 2.96\text{\AA}$, $|z_{an}| = 2$, $|z_{cat}| = 1$, $H/kT = 8$, and $N_{I2} = 1$.
- Figure 6. Distribution coefficient K_e as a function of hydrophobic interaction energy ϵ/kT for two values of M . Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = d_{cat} = 3.4\text{\AA}$, $|z_p| = 8$, $|z_{an}| = 1$, $|z_{cat}| = 1$, $H/kT = 6$, $I = 2.5$ moles/liter, $V = 0.006$, and $N_{I2} = 1$.
- Figure 7. Distribution coefficient K_e as a function of ionic strength I for three values of H/kT . Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = d_{cat} = 3.4\text{\AA}$, $|z_p| = 8$, $|z_{an}| = 1$, $|z_{cat}| = 1$, $H/kT = 6$, $\epsilon/kT = 4$, $M = 4$, $V = 0.006$, and $N_{I2} = 1$.
- Figure 8. Distribution coefficient K_e as a function of protein diameter d_p . Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = d_{cat} = 3.4\text{\AA}$, $|z_p| = 8$, $|z_{an}| = 1$, $|z_{cat}| = 1$, $H/kT = 6$, $I = 2.5$ moles/liter, $\epsilon/kT = 4$, $M = 4$, $V = 0.006$, and $N_{I2} = 1$.
- Figure 9. Distribution coefficient K_e as a function of ionic diameter for two ionic strengths. Other parameters are: $d_p = 40\text{\AA}$, $d_{an} = d_{cat} = d_{ion}$, $|z_p| = 8$, $|z_{an}| = 1$, $|z_{cat}| = 1$, $H/kT = 4$, $\epsilon/kT = 4$, $M = 4$, $V = 0.006$, and $N_{I2} = 1$.

Notation

a = Helmholtz energy, J/mol

a^{res} = residual Helmholtz energy, J/mol

$c_{ij}(r)$ = direct correlation function of i-j pair

C_s = salt concentration of solution, mol/L

d_i = diameter of a molecule, Å

e = elementary charge, 1.602×10^{-19} C

$g_{ij}(r)$ = radial distribution function of i-j pair

$h_{ij}(r)$ = total correlation function of i-j pair

H = Hamaker's constant, J

h = Planck's constant, 6.6252×10^{-34} J-sec

I = ionic strength of solution, mol/L

K_e = distribution coefficient

k = Boltzmann's constant, 1.3804×10^{-23} J/K

L_B = Bjerrum's length, Å

N_A = Avogadro's number, 6.023×10^{23} mol⁻¹

N_{I2} = ionic clustering parameter

m = mass of the molecule

M = number of hydrophobic sites per protein molecule

P = thermodynamic pressure, Pa

r = interparticle center-to-center distance, Å

T = absolute temperature, K

V = dimensionless association volume

$W(r)$ = potential of mean force, J

z = valence of ion

Z = compressibility factor

Greek Letters

$$\beta = 1/kT$$

ϵ_r = relative permittivity

ϵ_0 = permittivity in vacuum, C/Vm

ϵ = association energy

η = volume fraction

κ = inverse Debye length, \AA^{-1}

μ = chemical potential, J/mol

ρ = number concentration

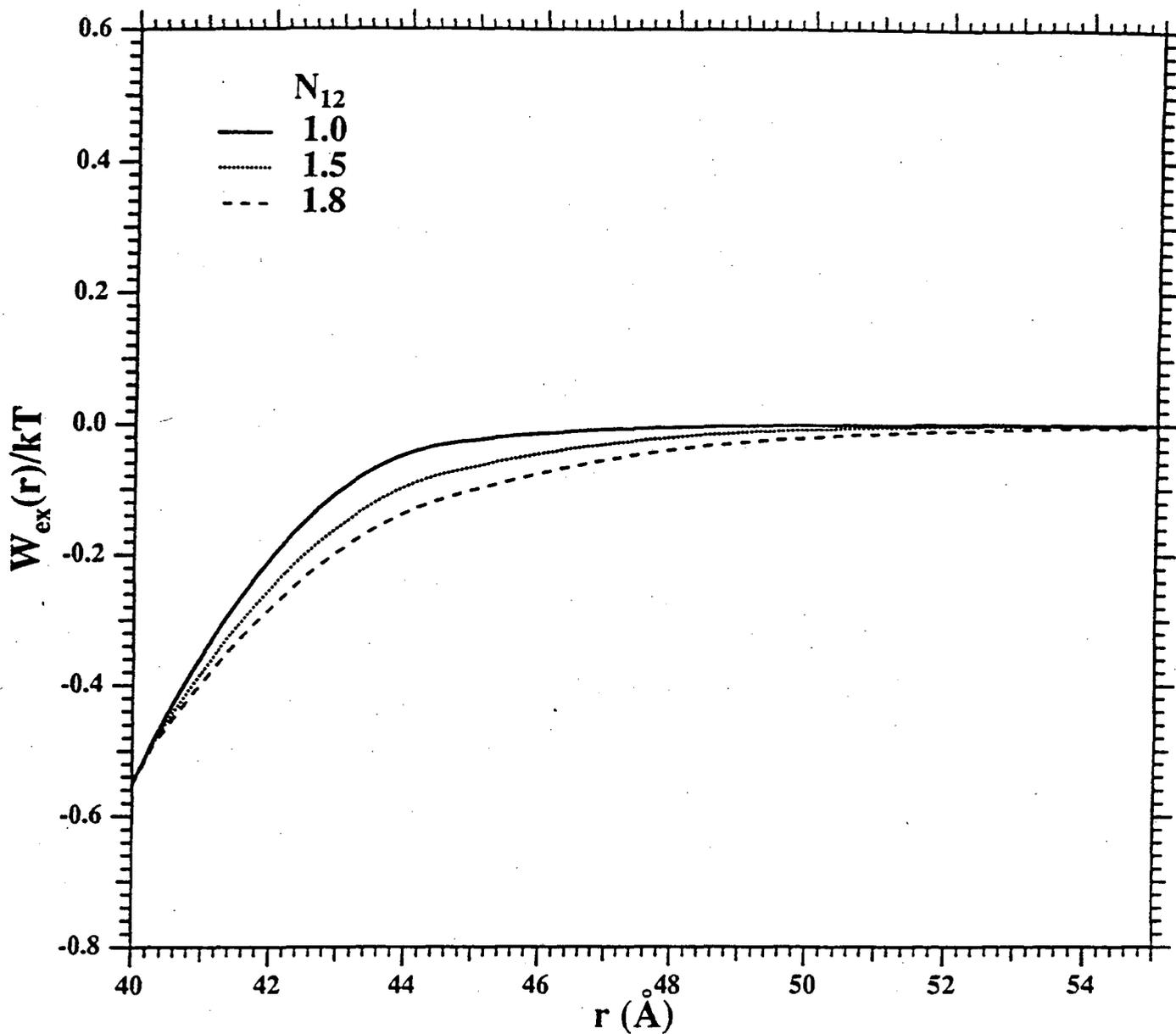
Subscripts

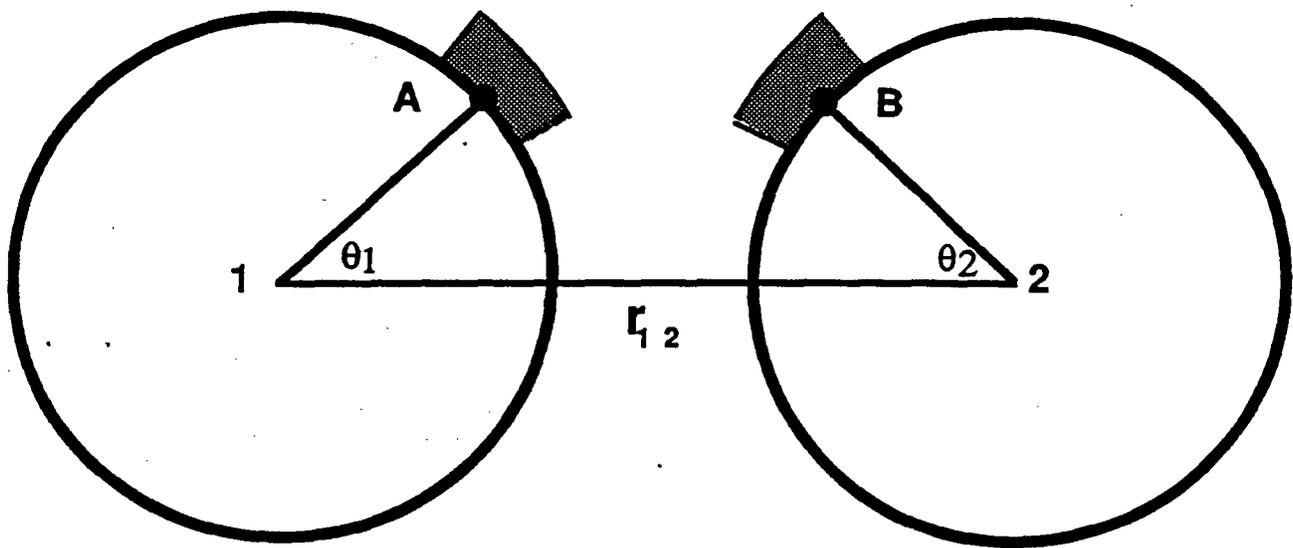
1 = ion

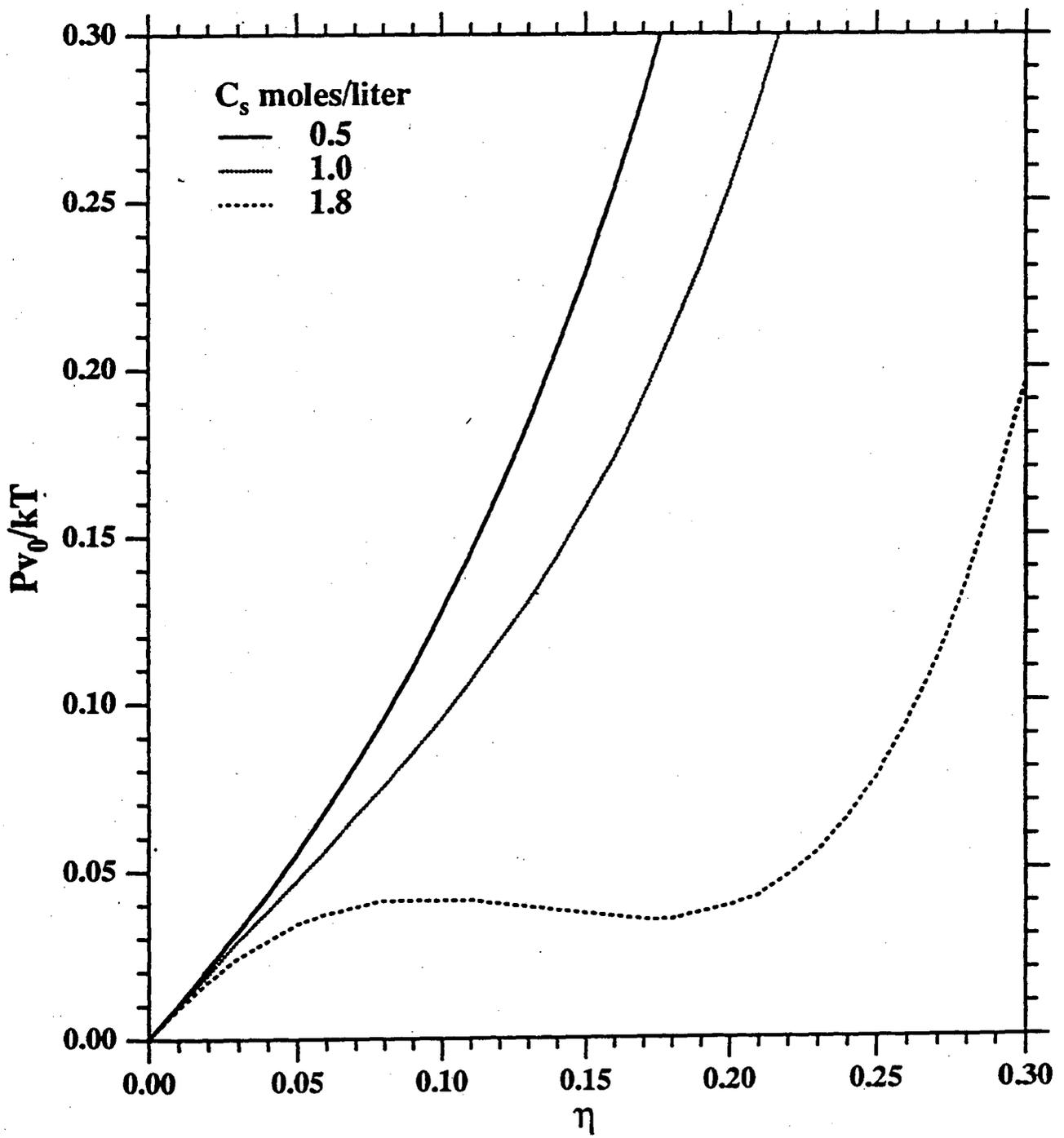
2 = ion

3 = protein

p = protein







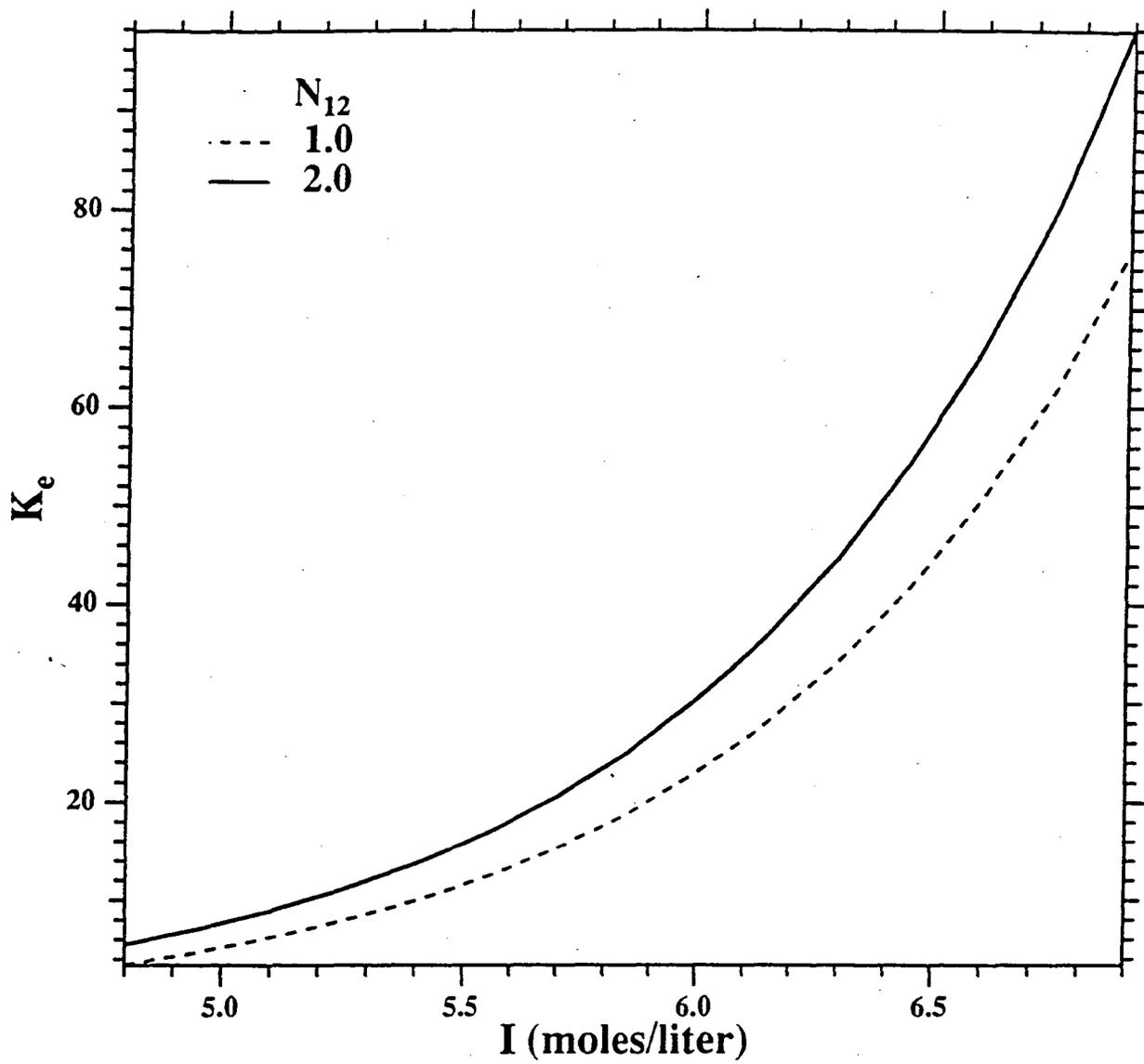


Figure 4

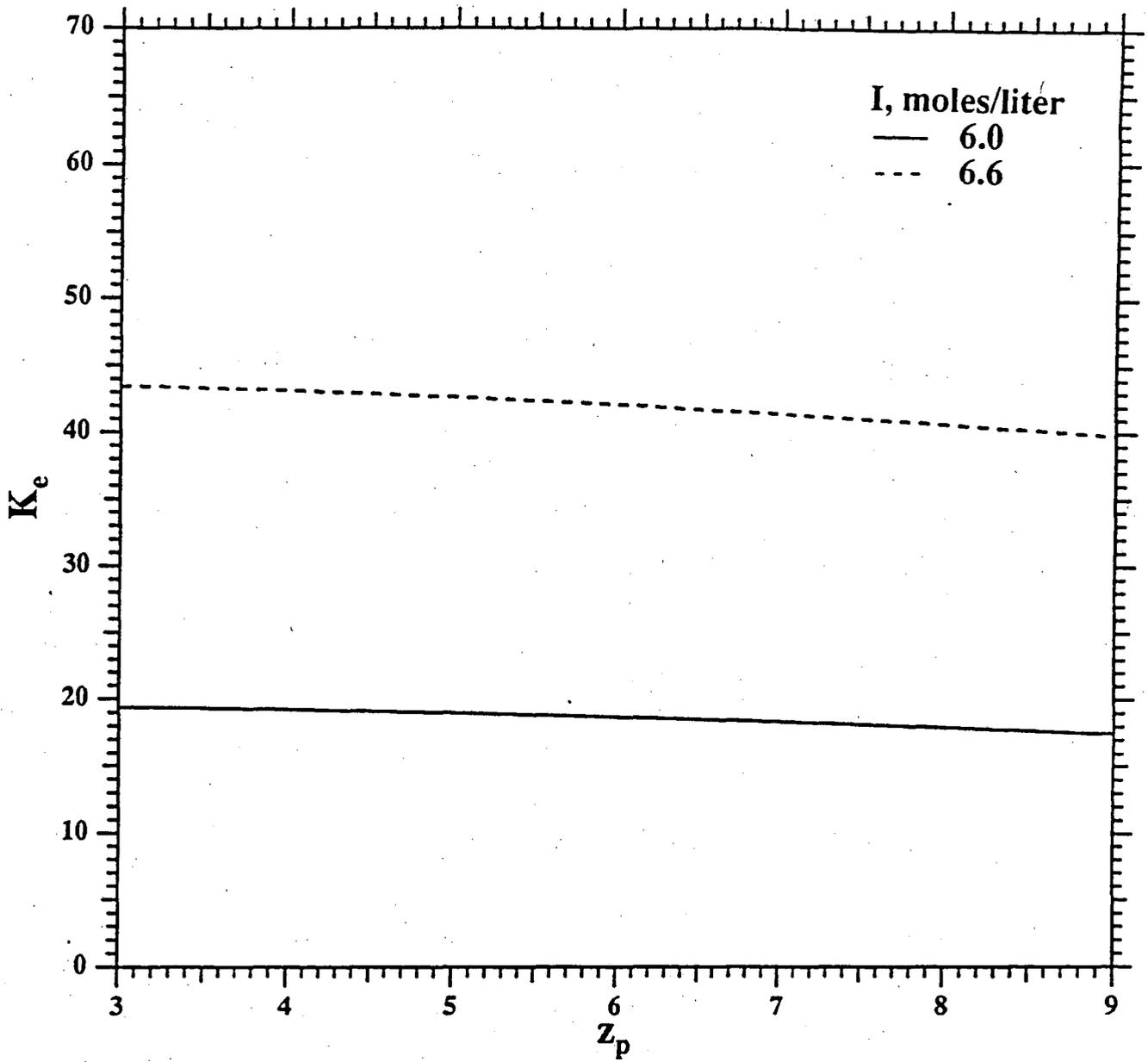
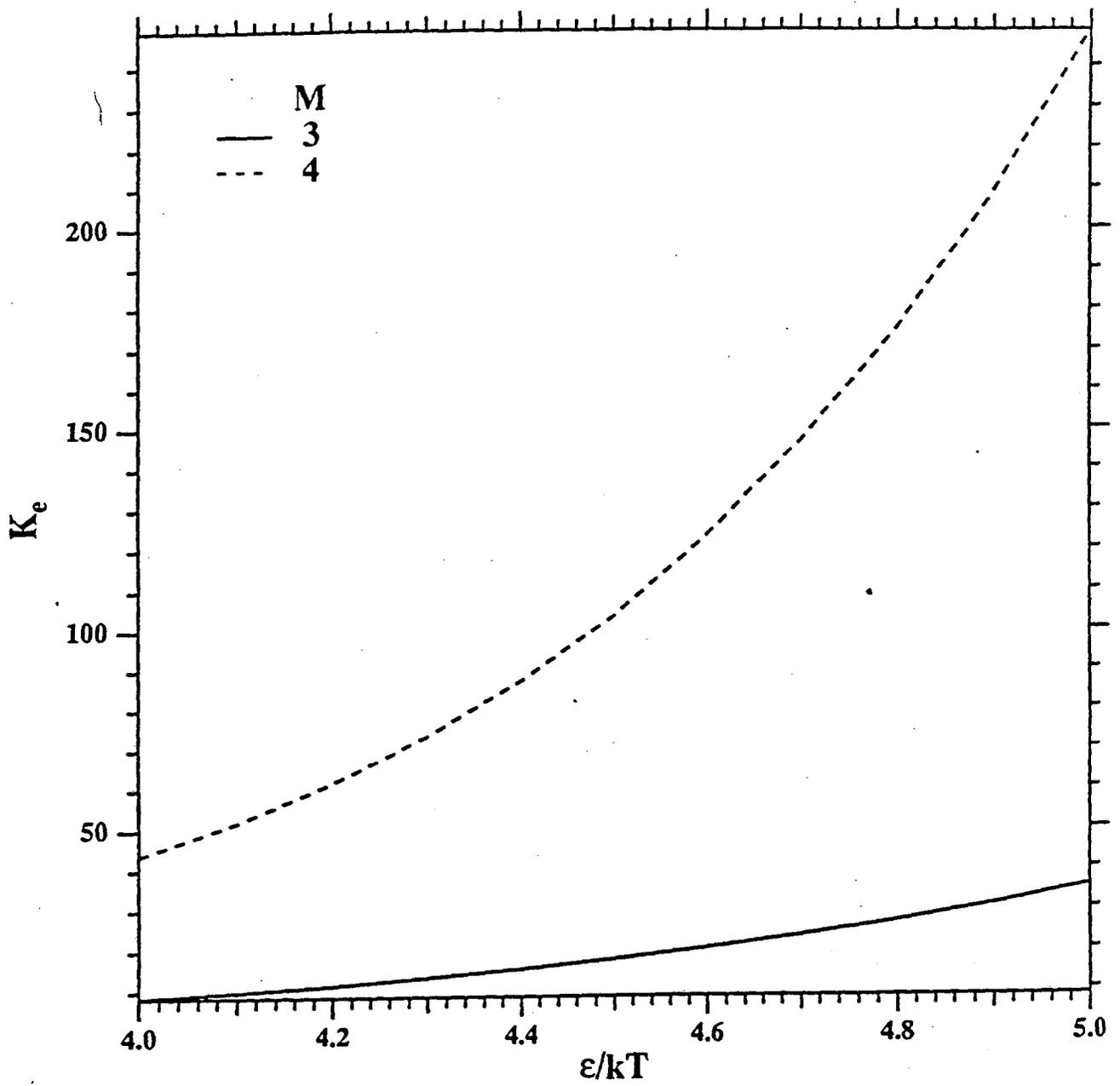


Figure 5

Chiew, Bland, Prausnitz



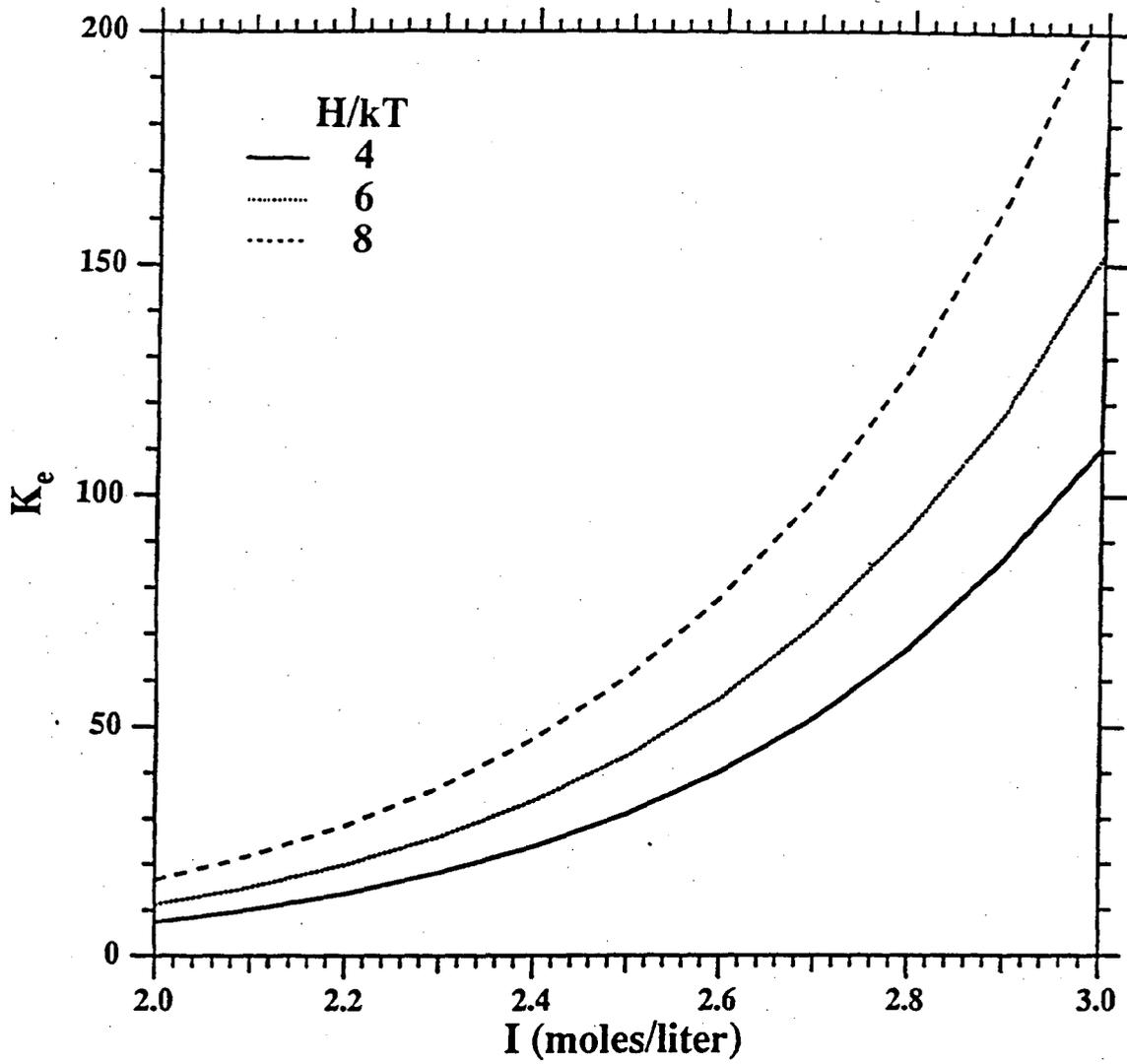


Figure 7

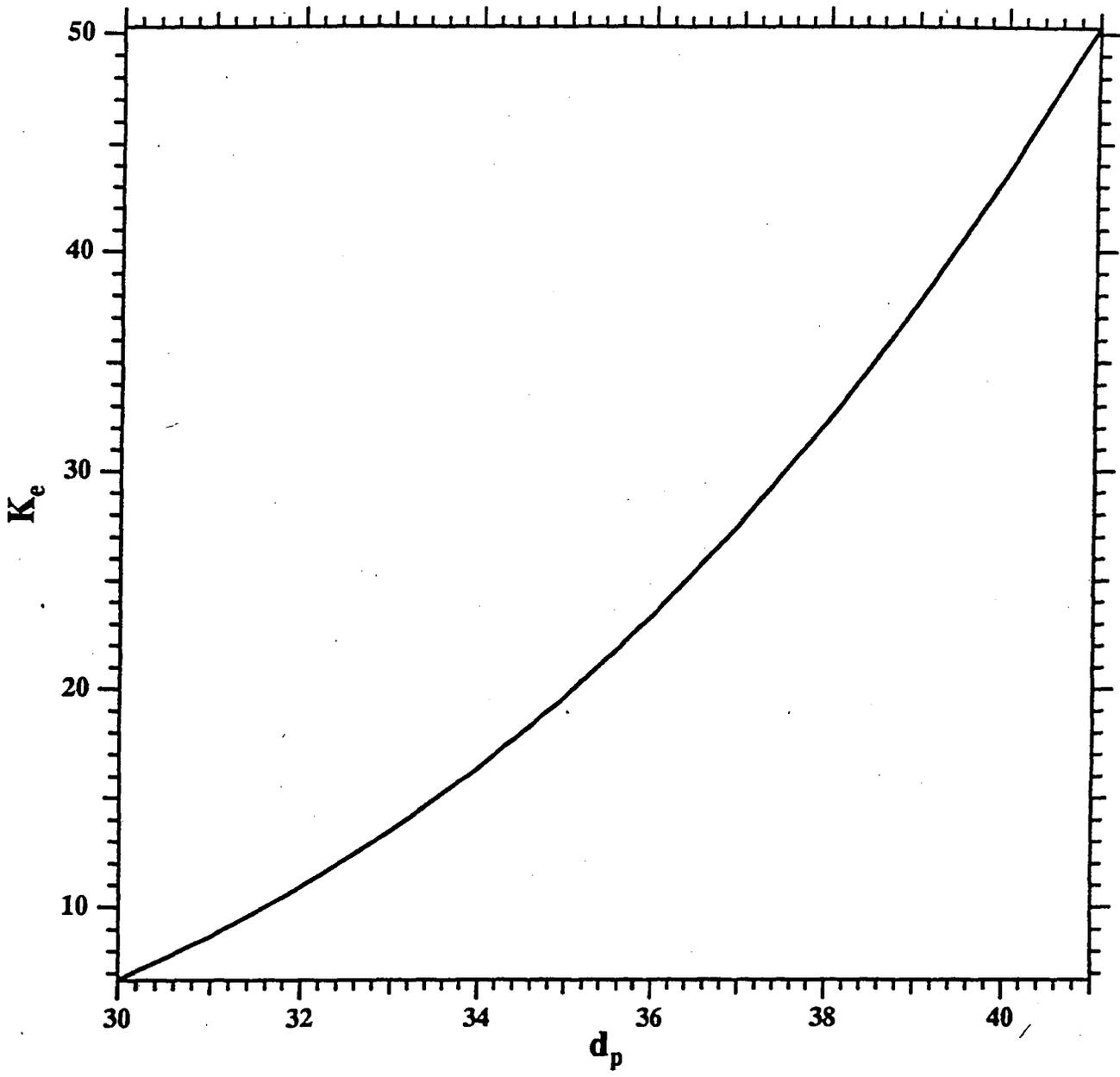


Figure 8

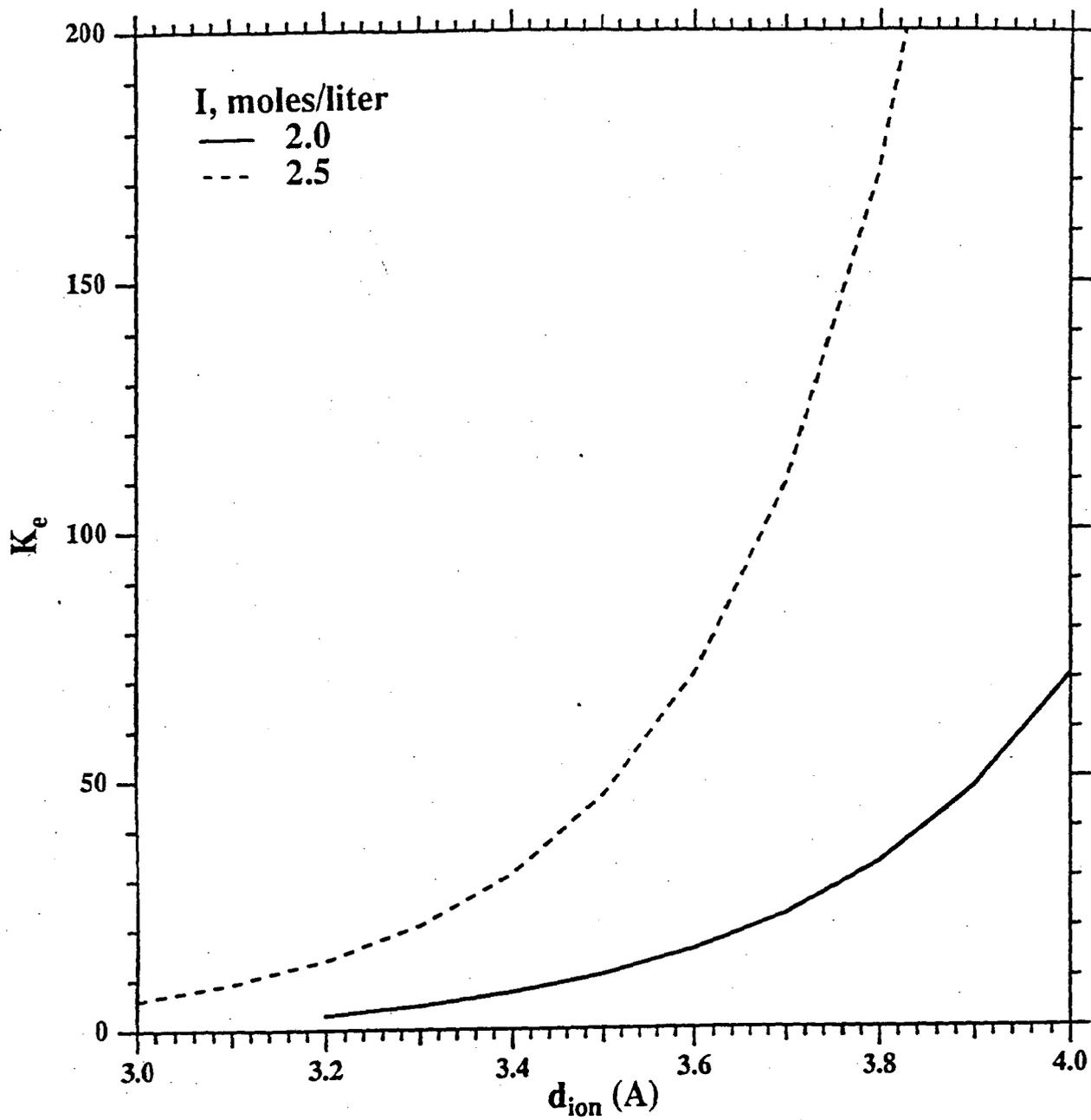


Figure 9

Chew, Blanch, Prausnitz

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720