Charge effect on protein partitioning in aqueous two-phase systems

Weiyu Fan
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Chemical Engineering Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
Charge effect on protein partitioning in aqueous two-phase systems

by

Weiyu Fan

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemical Engineering

Major Professor: Charles E. Glatz

Iowa State University

Ames, Iowa

1997
Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Weiyu Fan

has met the thesis requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College
# TABLE OF CONTENTS

**ABSTRACT** .................................................................................................................. vi

**CHAPTER 1. INTRODUCTION** ......................................................................................... 1

1. Dissertation Organization ............................................................................................. 3

**CHAPTER 2. LITERATURE REVIEW** ............................................................................. 4

1. Aqueous Two-Phase Systems ....................................................................................... 4

2. Protein Partitioning in Aqueous Two-Phase Systems ................................................. 10

3. Different Models Predicting Protein Partitioning Behavior in Aqueous Two-Phase Systems .................................................................................................................. 12
   a. Electrostatic effects and electrochemical partitioning .............................................. 12
   b. Electrostatic potential difference $\Delta \Phi$ and quasi-electrostatic potential difference $\Delta \Phi$ ........................................................................................................... 14
   c. Zaslavsky's view about electrochemical partitioning ............................................. 16
   d. Models accounting for non-electrostatic effects .................................................... 18

4. The Second Virial Coefficient and Laser Light Scattering ......................................... 23

**CHAPTER 3. CONTRIBUTION OF PROTEIN CHARGE TO PARTITION IN AQUEOUS TWO-PHASE SYSTEMS** ........................................................................... 28

**ABSTRACT** ...................................................................................................................... 28

**INTRODUCTION** ........................................................................................................... 29

**THERMODYNAMIC FRAMEWORK** ............................................................................ 31

**EXPERIMENTAL** .......................................................................................................... 36
CHAPTER 4. PROTEIN PARTITION IN AQUEOUS PEG-DEXTRAN TWO-PHASE SYSTEMS: EFFECTS OF DIFFERENT IONS AND IONIC STRENGTHS

ABSTRACT

INTRODUCTION

MATERIALS AND METHODS

Proteins

Aqueous two-phase systems

Protein partition coefficient measurements

Potential measurements

Flame emission adsorption measurements

RESULTS AND DISCUSSION

CONCLUSIONS

ACKNOWLEDGMENT

REFERENCES
Protein partitioning in aqueous two-phase systems is strongly affected by the net charge of the protein, but a thermodynamic description of the charge effects has been hindered by conflicting results. Many of the difficulties may arise from failures to isolate electrochemical effects from other interactions of phase components.

This work was targeted at characterizing the charge effects on protein partitioning in polymer-polymer aqueous two-phase systems by using two series of genetically-engineered charge modifications of bacteriophage T4 lysozyme produced in *E. coli*. The two series, one in the form of charged-fusion tails and the other in the form of charge-change point mutations, provided matching net charges but very different polarity. Partition coefficients of both series were obtained and interfacial potential differences of the phase systems were measured. Phase systems with different polymer concentrations, different salts and salt concentrations were investigated. T4 lysozymes with different surface charge display different partitioning behaviors in all the aqueous two-phase systems that were studied, which indicates the importance of electrostatic effect on protein partitioning. All the proteins have different partition coefficients in systems with different overall polymer concentrations because of different potential differences and different polymer-protein interactions between the two phases. T4 lysozyme fusion-tail mutants display different partitioning behavior in all the phase systems from those of point mutations with matching net charge. In the systems
with the same overall polymer concentration, but different salts and salt concentrations, both salts and proteins have different distributions between the two phases.

A model based on the virial expansion was established to describe the dependence of the protein partition coefficient on phase composition and potential difference. Multi-angle laser light scattering measurements were performed initially to determine the second virial coefficients between proteins and polymers in a phase solution to calculate the interaction parameter in the model. However, the prediction of the parameter from the second virial coefficients obtained from laser light scattering has been very difficult. Therefore, a study on the second virial coefficients of different T4 lysozyme mutants in solutions with different ionic strengths was performed. The values of the second virial coefficients obtained from laser light scattering measurements were compared to those calculated from McMillan-Mayer theory using estimates of the molecular interaction forces. The results from laser light scattering agreed with the theoretical calculations at higher but not lower ionic strengths. Comparison of the second virial coefficients between T4 lysozymes with different net charge or charge distribution reveals the effects of surface charge and dipole moment of the protein on electrostatic intermolecular interactions.
CHAPTER 1. INTRODUCTION

First described by Beinerinck in 1896, aqueous two phase systems usually contain a water solution of two polymers such as poly (ethylene glycol) and dextran, or one polymer and certain salts. Spontaneous phase separation occurs when the concentration of polymers or salts increases to about 5-10% weight concentration.

Aqueous two phase systems were first used for bioseparation by Albertsson (Albertsson, 1956). In the 80's, with the development of recombinant DNA technology and protein engineering, it has become possible to produce a large number of clinically useful protein products and bulk industrial enzymes by microorganisms and this brings a demand for a new low cost and high efficiency separation method. Since then, more and more research on phase systems and their applications have been conducted for several reasons. First, as a water solution, a phase system provides mild and nondeleterious conditions for most biomaterials. Secondly, the selection of cheap polymers and the short processing time required for phase equilibrium make the aqueous two-phase separation a potentially low cost process. Finally, the liquid-liquid extraction is easy to scale-up (Albertsson, 1986. Walter et al., 1985, Zaslavsky, 1995).

Considerable advances in the theory of two-phase systems and partitioning phenomena have been made(Albertsson, 1986, Walter et al., 1991, Abbot et al., 1990. Walter et al., 1985, Zaslavsky, 1995). Different properties of aqueous two-phase systems have been characterized and many models have been proposed to predict the phase diagrams
and partitioning coefficients. However, due to the complexity of the electrolyte and polymer solution structure, most of these models reflect only one or two aspects of phase separation mechanisms so the application of these models is still limited.

In this work, two series of genetically engineered charge modifications of T4 lysozyme are employed to study charge effects on the partitioning behavior. The specific objectives of this work were:

1. To study the effect of protein net charge and charge distribution on the protein partitioning in aqueous two-phase systems.

2. To study the electrostatic potential difference between the two phases of a phase system and its effect on the charged protein partitioning in phase systems.

3. To investigate how charge-change mutations can bring changes in protein-protein and protein-polymer interactions in phase solutions and how these changes affect the protein partitioning in aqueous-two phase systems.

Protein partitioning in PEG/dextran aqueous two-phase systems with different polymer concentrations, different salts and salt concentrations were determined. The potential differences between the two phases were measured. Multi-angle laser light scattering was used to characterize the different molecular interactions in the phase solution. A model based on the virial expansion was established to predict the protein partition coefficient. These results not only provide theoretical insights of protein partitioning in aqueous two-phase systems, but also enable us to combine manipulation of protein charge, polymer concentration, and salt content of the phase system to enhance extraction selectively.
1. Dissertation.

The dissertation provides a gene...next section consists with the protein partition interacts in...Chapter 6. Ref...
DISSERTATION ORGANIZATION

The dissertation contains three main sections. The first section (Chapter 1.2) provides a general introduction to some of the basic principles and concepts of aqueous two-phase systems followed by a comprehensive review of the pertinent literature. The next section consists of three journal manuscripts (Chapters 3-5). Chapters 3 and 4 deal with the protein partitioning in different aqueous two-phase systems and chapter 5 investigates the second virial coefficient of different genetically-engineered proteins and their molecular interactions in protein solutions. The dissertation concludes with the overall conclusions in Chapter 6. References cited in Chapters 1-2 are listed at the end.
CHAPTER 2. LITERATURE REVIEW

1. Aqueous Two-Phase Systems

Pairs of polymers or one polymer and one salt can display incompatibility when dissolved in a water solution above certain concentrations. Possible compositions of phase systems are listed in Table 2.1. Essentially all these non-ionic water-soluble polymers can be viewed as composed of two kinds of groups. One kind is hydrophobic groups and the other is groups that can form H-bonding or dipole-dipole interactions. For example, poly(ethylene glycol) (PEG) has a hydrophobic region (-CH2-CH2-) and an H-bonding ether oxygen per monomer unit (Albertsson, 1986, Brooks, 1985, Zaslavsky, 1995).

Obviously water plays an important role in the forming of phase systems. Water can interact with and bind to polymer molecules by hydrogen-binding or other hydration effects. Albertsson (1986) proposed a hydrophobic ladder of phase polymers and found that the hydrophobicity of an aqueous polymer solution is very close to that of water or salt water solution (Figure 2.1). Zaslavsky (Zaslavsky 1995) also compared the polarity of aqueous polymer solutions to water and a similar conclusion was reported (Figure 2.2).

According to the second law of thermodynamics, two components will mix if the Gibbs free energy of mixing, $\Delta G_{mix} < 0$,

$$\Delta G_{mix} = \Delta H_{mix} - T \cdot \Delta S_{mix} < 0$$

(2.1)

where $\Delta H_{mix}$, $\Delta S_{mix}$ are the enthalpy and entropy of mixing, and $T$ is the temperature.
<table>
<thead>
<tr>
<th>A. Polymer-polymer-water systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>poly(ethylene glycol)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>poly(Propylene glycol)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>poly(vinyl alcohol)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>poly(vinyl pyrrolidone)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>dextran</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Polymer-low molecular weight solute-water systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>poly(ethylene glycol)</strong></td>
</tr>
<tr>
<td><strong>poly(propylene glycol)</strong></td>
</tr>
<tr>
<td><strong>poly(vinyl pyrrolidone)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>poly(propylene glycol)</strong></td>
</tr>
<tr>
<td><strong>dextran</strong></td>
</tr>
</tbody>
</table>
### Figure 2.1 Hydrophobic ladder.

The hydrophobicity of an aqueous solution of polymers is very close to that of water or salt water.

<table>
<thead>
<tr>
<th>Polymer</th>
<th>Added to Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenol</td>
<td>Polypropylene glycol + H₂O</td>
</tr>
<tr>
<td>Salt + H₂O</td>
<td>Polyethylene glycol + H₂O</td>
</tr>
<tr>
<td>Ether</td>
<td>Polyvinylalcohol + H₂O</td>
</tr>
<tr>
<td>Phenol</td>
<td>Methylcellulose + H₂O</td>
</tr>
<tr>
<td>Acetone</td>
<td>Hydroxypropyldextran + H₂O</td>
</tr>
<tr>
<td>H₂O</td>
<td>Dextran + H₂O</td>
</tr>
<tr>
<td>Salt + H₂O</td>
<td>Carboxymethyl dextran + H₂O</td>
</tr>
<tr>
<td></td>
<td>Dextran sulfate + H₂O</td>
</tr>
</tbody>
</table>

### Figure 2.2 Polarity ladder.

Relative solvent polarity of polymer solutions are close to that of water.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Polymer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diethyl ether</td>
<td>PEG-6000</td>
</tr>
<tr>
<td>Chloroform</td>
<td>PEG-200</td>
</tr>
<tr>
<td>Cyclohexanol</td>
<td>Ficoll-400</td>
</tr>
<tr>
<td>1-Octanol</td>
<td>Dextran-70</td>
</tr>
<tr>
<td>1-Butanol</td>
<td></td>
</tr>
<tr>
<td>Ethanol</td>
<td></td>
</tr>
<tr>
<td>Methanol</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td></td>
</tr>
</tbody>
</table>
Since enthalpy indicates the molecular interaction energy and entropy is a measure of disorder. Equation 2.1 indicates that phase separation takes place when the entropy increase, i.e. the increase of disorder from mixing of the two polymers, is overcome by the increase of enthalpy upon mixing, i.e. the loss of favorable molecular interactions during mixing. To calculate the enthalpy and entropy of mixing, however, we will have to consider the non-idealities of the solution.

A phase diagram describes the partitioning behavior (concentrations of the two polymers in each phase) of an aqueous two-phase system (Figure 2.3). Curve TCB is called a binodal curve and represents the polymer concentrations in the two phases at equilibrium. Above the binodal curve is the two-phase region where phase splitting occurs and below the curve is the one-phase region where an aqueous solution remains homogenous. In Figure 2.3, system A will split into top phase T and bottom phase B. TB is the tie line and the length of TB is the tie line length (TLL), which indicates the concentration difference of the two polymers between top phase and bottom phase:

\[ TLL = \left( \Delta C_{\text{Polymer}_2}^2 + \Delta C_{\text{Polymer}_3}^2 \right)^{1/2} \]  \hspace{1cm} (2.2)

where as a conventional way of expressing phase components, subscript 1 denotes solvent, i.e. water, and 2 and 3 denote the two polymers.

From conservation of mass, the ratio of phase volumes \( R \):

\[ R = \frac{V_{\text{top}}}{V_{\text{bottom}}} = \frac{AB}{TA} \]  \hspace{1cm} (2.3)

Because of the availability, wide ranges of molecular weight, and convenient viscosities of the two polymers (Albertsson, 1986), PEG-dextran aqueous two-phase systems
Figure 2.3 A general phase diagram for a two-phase system. Tie line TAB passes through the total composition, A, and phase compositions, T and B; curve TCB is the binodal curve and C is the critical point.

have been widely investigated and used for different separation purposes. Many results of protein partitioning in PEG-dextran systems have been published (Albertsson, 1986; Zaslavsky, 1995). Therefore, we choose PEG-dextran system in our research work. Figure 2.4 shows the structures of PEG and dextran. The molecular weights of PEG and dextran typically range from 1,000 - 40,000 and from 40,000 - 2,000,000 respectively.

Phase diagrams of the PEG and dextran system under different conditions have been reported (Albertsson, 1986; Zaslavsky, 1995). The position of the binodal curve of PEG-
Figure 2.4 Molecular Structure of PEG (a) and Dextran (b)

dextran phase diagram is affected by the molecular weight of the polymers, temperature and the presence of different inorganic salts. Higher molecular weight polymers lead to lower polymer concentrations for phase separation and a higher temperature will require a higher concentration for phase separation. The influence of addition of salts is somewhat complicated. Effectiveness in promoting phase separation have been reported as (Abbott. 1990):

$$\begin{align*}
PO_4^{3-} & > SO_4^{2-} > F^- > Cl^- > Br^- > I^- \\
K^+ & = Na^+ > Cs^+ > Sr^+ > Ca^{2+} > NH_4^+ > Li^+ 
\end{align*}$$

(2.4)
However, others (Cabezas et al., 1990; Walter, 1991) have claimed that sulfite has a dramatic effect on the PEG-dextran phase separation. The influence of salts reflects the fact that solvent structure is very important in phase formation. Qualitative explanations include salt effects on water structure, charge-charge and charge-dipole interactions between polymers and salts, and the competition between polymers and salts for water of hydration (Zaslavsky, 1995; Zaslavsky et al., 1987; Albertsson, 1986).

2. Protein Partitioning in Aqueous Two-Phase Systems

Proteins usually distribute unevenly between the two phases in an aqueous two-phase system and their partitioning between the two phases depends on many factors. A partition coefficient $K$ is used to characterize the protein partitioning behavior:

$$K = C_T C_B$$

where $C_T$ and $C_B$ are the protein concentrations in the top and bottom phase respectively.

The partition coefficient can be expressed as a product of different terms, each representing a kind of physical property or interaction contributing to the observed value (Albertsson, 1985):

$$\ln K = \ln K_o + \ln K_{el} + \ln K_{hpho} + \ln K_{biosp} + \ln K_{size} + \ln K_{conf}$$

(2.6)

where el, hpho, biosp, size, and conf stand for electrochemical, hydrophobic, biospecific, size and conformational contributions to the partition coefficient and $\ln K_o$ includes all other factors. Details of models for $K_o$ and electrochemical partitioning will be discussed in the next chapter, while the others are reviewed here.
Hydrophobic affinity refers to separation in the phase system based on the hydrophobicity of protein molecules. It is found that in the PEG-dextran system, solutes that are hydrophobic tend to partition preferentially into the upper (i.e. PEG-enriched) phase. Eiteman et al. (1994) reported the calculation results of partition coefficients of peptides from the hydrophobicities of the individual amino acids; and the partition coefficient of protein can be enhanced by inserting an hydrophobic peptide handle into the protein.

Biospecific affinity is the interaction between sites on the desired protein molecules and ligands attached to the phase polymers. Reported ligands include metals, triazine dyes, charged polymers, etc. (Arnold, 1991; Cheng, et al., 1990; Cordes, et al., 1987). The presence of ligands can increase the selectivity and yield of protein partitioning. The increase is usually expressed as $\Delta \ln K_p$,

$$\Delta \ln K_p = \ln \left( \frac{K_{p,\text{affi}}}{K_p} \right)$$

where $K_{p,\text{affi}}$ is the protein partition coefficient measured in the presence of the ligand-bound polymer and $K_p$ is the partition coefficient in the absence of ligands.

Size and conformation of a protein can also affect its partitioning. It is generally observed that as the size of the protein or peptide increases, the distribution is more uneven between the two phases (Sasakawa and Walter, 1972). Small molecules such as dipeptides usually have a partition coefficient close to 1 (Diamond and Hsu, 1989). However, it is hard to decide whether the partition coefficient depends on the volume of the molecule or the surface area of the molecule.
Equation 2.6 splits the partitioning coefficient $K$ into several terms and it seems applicable when one term dominates the others. But in most aqueous two-phase systems, all the above factors will affect the partitioning and they are not independent of each other. Moreover, not every term in the model can be strictly related to a certain type of interaction force between molecules.

3. Different Models Predicting Protein Partitioning Behavior in Aqueous Two-Phase Systems

a. Electrostatic effects and electrochemical partitioning

The concept of electrochemical partitioning is based on the interfacial potential difference between the two phases of a phase system caused by the Donnan effect due to different solubility or distribution of ionic species between two phases (Johansson, 1974a, b, 1985; Albertsson, 1986; Luther and Glatz, 1994; Schluck, 1995a, b; Großmann et al., 1995; Forciniti et al., 1990, 1992). Because of the existence of this potential difference, the potential of an ionic species in a phase solution is actually an electrochemical potential and can be divided into chemical and electrical contributions (Newman, 1991; Haynes et al., 1991):

$$\tilde{\mu}_i = \mu_{i,\text{chem}} + z_i F \phi$$  \hspace{1cm} (2.8)
where \( \tilde{\mu}_i \) is the electrochemical potential of species \( i \); \( \mu^\text{chem}_i \) is the chemical potential of species \( i \); \( z_i \) is the charge of the ionic species and \( \varphi \) is the "purely" electrostatic potential.

Hence, the usual formulation

\[
\tilde{\mu}_i = \mu^\text{chem}_i + RT \ln m_i \gamma_i
\]  

becomes

\[
\tilde{\mu}_i = \mu^\text{chem}_i + RT \ln m_i \Gamma_i + z_i F \varphi
\]  

(2.10)

where \( \mu^\text{chem}_i \) is the chemical potential at the standard state; \( \gamma_i \) is the activity coefficient; \( m_i \) is the concentration; \( z_i \) is the charge of the ionic species; \( \Gamma_i \) is a "chemical" activity coefficient and \( \varphi \) is the "purely" electrostatic potential. At the phase equilibrium, we can have:

\[
\tilde{\mu}_i^T = \tilde{\mu}_i^B
\]  

(2.11)

where superscripts \( T \) and \( B \) denote top phase and bottom phase respectively. Thus, the partition coefficient of species \( i \) can be expressed by substituting Equation 2.10 into Equation 2.11 for both top and bottom phase:

\[
\ln K_i = \ln \frac{m_i^T}{m_i^B} = \ln \frac{\Gamma_i^T}{\Gamma_i^B} + \frac{z_i F}{RT} (\varphi^B - \varphi^T)
\]  

(2.12)

so \( K_i \) is seen to account for chemical effects caused by non-electrostatic interactions and \( \Delta \varphi = \varphi^B - \varphi^T \), is the electrostatic potential difference.

Equation 2.12 was derived (Albertsson, 1986) as a model predicting the protein electrochemical partitioning, where the electrostatic force dominates the partitioning process.

This model assumes the protein has a point charge equal to its net charge and the protein
concentration is assumed to be sufficiently low so that it won't disturb the potential difference between the two phases. Many studies have been conducted in testing and applying Equation 2.12 and different results were obtained (Johansson, 1970, 1974a, b, 1985; Albertsson, 1986; Abbott et al., 1990). More often, the potential difference has been calculated from the slope of a plot of measured $\ln K_p$ versus $z_p$. Using this method, while varying $z_p$ by changing pH, potential differences in the range of 0-20 mV were found (Forciniti et al., 1992, 1990; Hartounian et al., 1994; Haynes et al., 1989; Johansson, 1974a, 1985; Luther and Glatz, 1994, 1995; Schluck et al., 1995a, b). However, a linear relationship between $z_p$ and $\ln K_p$ was not always obtained. Nonlinearity can result from pH dependence of $\Delta \phi$, protein conformation and phase compositions. (Forciniti et al., 1992; Schluck et al., 1995a, b; Zaslavsky, 1995). The latter changes would directly alter $K_o$.

b. Electrostatic potential difference $\Delta \phi$ and quasi-electrostatic potential difference $\Delta \Phi$

The potential differences measured directly by using Ag / AgCl electrodes have also been reported (Johansson, 1974a, 1985; King et al., 1988; Haynes et al., 1991, 1989; Luther and Glatz, 1994). However, $\Delta \phi$ in Equation 2.12, the electrostatic potential difference, according to Guggenheim (1959), is not measurable. The potential difference measured with Ag/AgCl electrodes, $\Delta \Phi$, is different from the pure potential difference and is actually the difference of quasi-electrostatic potential of the reference ion $r$ (Newman, 1991; Haynes et al., 1991). The quasi-electrostatic potential, $\Phi$, is defined as:
where $\mu_r$ is the electrochemical potential of the reference ion $r$; $\mu_r^0$ is the chemical potential at the standard state; $\gamma_r$ is the activity coefficient and $m_r$ is the molality of the reference ion.

This could explain the discrepancy between the measured potential difference and the potential difference calculated from the Equation 1 using protein partitioning data (Luther and Glatz, 1994, 1995). The demonstrated usefulness of Equation 1's treatment of electrostatic effects (Johansson, 1974a, 1985; Hartounian et al., 1994; King et al., 1988; Luther and Glatz, 1994; Schluck et al., 1995a, b) should then be attributed to an empirical accounting for activity coefficient effects. To see the connection between $\Delta \varphi$ and $\Delta \Phi$ (Haynes et al., 1991; Newman, 1991), consider that $\Delta \Phi$ is measured experimentally by sensing a reference ion (such as the $K^+$ ion in a salt bridge between the two phases). From Equation 2.13, we have:

$$ z_r F \Phi = \mu_r^0 + RT \ln \gamma_r \quad (2.14) $$

and

$$ \Delta \Phi = \Phi^b - \Phi^T = -\frac{RT}{z_r F} ln \frac{m_r^b}{m_r^T} \quad (2.15) $$

with \( \frac{m_r^T}{m_r^b} = K_r \). Hence, the measured potential difference, $\Delta \Phi$, is properly linked to partitioning of the reference ion (assuming junction potential errors are negligible).

When turning to the other charged species (such as Equation 1 is to be applied to), one can combine Equation 2.9 with Equation 2.14 and giving

$$ \tilde{\mu}_r = \mu_r^0 + RT \ln m_r + z_r F \Phi - \frac{z_r}{z_r} (\mu_r^0 + RT \ln \gamma_r) \quad (2.16) $$
Writing Equation 21 for both phases at equilibrium, we find

\[ \ln K_i = \ln \frac{m_i^r}{m_i^s} = \frac{z_i F}{RT} (\Phi^s - \Phi^r) + \frac{z_i}{z_r} \ln \frac{\gamma_r^T / \gamma_r^s}{(\gamma_r^T / \gamma_r^s)^{z_r/z_i}} \]  

(2.17)

which shows that, as used,

\[ \ln K_0 = \frac{z_i}{z_r} \ln \frac{\gamma_r^T / \gamma_r^s}{(\gamma_r^T / \gamma_r^s)^{z_r/z_i}} \]  

(2.18)

Therefore, when the measured potential difference, \( \Delta \Phi \), is substituted into Equation 1, \( K_0 \) is also related to the activity of the reference ion in the two phases.

c. Zaslavsky's view about electrochemical partitioning

On the other hand. Zaslavsky argued that treating the electrochemical partitioning in aqueous two-phase systems by using Equation 2.12 is not proper (Zaslavsky, 1995). First of all, it is impossible to study a single ion solution, e.g. a solution of Na\(^+\), and to measure a single ion thermodynamic characteristic, such as ionic solvation free energy, enthalpy, partition coefficient, etc., but in Equation 2.12, protein is treated as a single ionic species. In fact, because of the electroneutrality constraints, transfer of protein from one phase to another must be accompanied by the transfer of the electrically equivalent amount of the corresponding counter-ions. So the partition coefficient value for an ionic solute such as protein actually characterizes partition behavior not of a single ion but of both the protein and its counter ions. Moreover, for a \( B_x A_y \), we have:

\[ x \cdot z^+_b = -y \cdot z^-_a \]  

(2.19)

and
\[ x \cdot z_B^- \cdot F \cdot \Delta \varphi = -y \cdot z_A^+ \cdot F \cdot (-\Delta \varphi) \]  

(2.20)

which means the electrical energy contributions in the partition coefficient as indicated by the second term in Equation 2.12 are canceled out.

Therefore, Zaslavsky argues that 'an interfacial potential difference does not affect partitioning of ionic solutes because the electrical interfacial work terms are canceled out': and the partitioning is a result of different intensity of the electrostatic ion-ion and ion-dipole interactions in the two phases because of the different dielectric properties and polymer concentrations of the two phases.

Very few efforts on reconciling the two controversial arguments have been reported. Some believe these two viewpoints are not inconsistent as the intermolecular forces are the origin of the interfacial potential difference (Abbott et al., 1990). Zaslavsky's argument appears a framework more based on the molecular interactions in the solution. Using this perspective, Neogi (1993) employed Hill's (1964) small system thermodynamics theory to describe the electrostatic effect based on the molecular interactions. The role of charge distribution around a protein molecule was systematically analyzed. The model indicates that the bulk potential difference doesn’t affect the charge distribution around the protein molecule so that \( \Delta \varphi \) has no impact on the protein partition coefficient. Taking the upper limit to the electrostatic effects on \( K_p \) gives an expression that \( \ln K_p \) is proportional to the square of the protein net charge. The author showed \( \ln K_p \) from other people's experiments had a linear relationship with the square of the protein net charge; however, the prediction of the partitioning coefficient can not be attained.
d. Models accounting for non-electrostatic effects

As we have seen in Equation 2.12, the expression for protein partition coefficient is divided into electrostatic effects of $\Delta \phi$ term and non-electrostatic effects in $K_p$. For the latter, many models based on molecular interactions in phase solution have been developed (Abbott et al., 1990; Zaslavsky, 1995). Among them, there are two classes of models that have been studied and applied widely: one employs the Flory-Huggins theory (Flory, 1953; Huggins, 1942) and the other uses the osmotic virial expansion of Edmond and Ogston (1968).

The Flory-Huggins model, also called the lattice model, regards the liquid structure as a lattice similar to that of a crystal. A solvent molecule takes one site in the lattice. A polymer molecule can be cut into $r$ segments (e.g. a CH$_2$- in a polymer chain usually can be considered a segment) and each segment also occupies one site. Thermodynamic properties of the system then can be derived from statistical mechanics. Brooks et al. (1985) first applied this model to protein partitioning in aqueous two-phase systems by treating the protein as a third polymer component. For a system containing four components: 1-solvent, i.e. water; 2-polymer 1; 3-polymer 2; p-protein, the equilibrium protein partition coefficient can be expressed as:

$$
\ln K_p = r_p \left[ (\phi_1^p - \phi_1^s)(1 - \chi_{1p}) + (\phi_2^p - \phi_2^s)(1 - \chi_{2p}) \left( \frac{1}{r_2} - \chi_{3p} \right) \right] + (\phi_3^p - \phi_3^s)(1 - \chi_{3p}) \left( \frac{1}{r_3} - \chi_{3p} \right) 
$$

(2.21)
where \( r_i \) is the number of segments of the polymer molecule and also the number of sites occupied by species \( i \); \( \phi_i^t \) and \( \phi_i^b \) are the volume fraction of species \( i \) in the top and bottom phase respectively, \( \chi_{ij} \) is the so-called Flory-Huggins interaction parameter, depicting the energy change associated with the formation of a contact between the segments of polymer \( i \) and \( j \).

Since most globular proteins are dense, compact, and relatively rigid macromolecules, the treatment of the protein species as a third polymer, i.e. a diffuse random coil in the model raises some uncertainty in the results, so that the model can only describe the protein partitioning behavior in phase systems qualitatively. Baskir et al. (1987) tried to improve the lattice model by approximating the protein molecule as a rigid, impenetrable and spherical body of known size and with homogeneous surface properties, rather than a random-coil polymer species. A new parameter, \( \chi_s \) was used to characterize the free-energy change caused by the polymer segment-protein interaction in his expression. This model physically appears more reasonable and it allows the prediction of protein partitioning coefficient as well as the binodal curve. But the computation procedure is very complicated and the interaction parameters are too difficult to obtain. Different forms of simplification also have been reported (Diamond et al., 1989, 1990). Diamond et al. (1990) compared the magnitude of the five terms in the Flory-Huggins model for a PEG-Dextran phase system with low salt concentration, and were able to simplify the expression to two terms plus an electrostatic effect term:
where $k_1$ and $k_2$ contain the Flory-Huggins interaction parameters, the molar volume ratios, and the protein molecular weight; they can be considered constant for a specific protein in a specific phase system. $w_f^b$ and $w_f^T$ are weight fractions of polymer 2, i.e. PEG, in the bottom and top phases, respectively. The third term was added to the model to take into account the electrostatic effect in the same manner as Equation 2.12.

The virial expansion model was first applied to an aqueous two-phase system by Ogston and Edmond (1968). It is based on a mathematical expansion of the thermodynamic properties of the polymer solution in terms of the concentration of the polymers. King et al. (1988) applied the virial expansion model to derive the $K_p$ term in Equation 2.12 and expressed the protein partition coefficient as:

$$\ln K_p = a_{2p} (m_2^b - m_2^T) + a_{3p} (m_3^b - m_3^T) + \frac{z_p F}{RT} \Delta \phi$$

(2.23)

where $m_2$ and $m_3$, defined as mole solute/kg, are the molar concentration of polymer 2 (PEG) and polymer 3 (dextran) in phase systems, and can be related to the weight fraction concentrations $w_2$ and $w_3$ as:

$$m_2 = \frac{1000 w_2}{M_2} \frac{1}{1 - w_2 - w_3}$$

$$m_3 = \frac{1000 w_3}{M_3} \frac{1}{1 - w_2 - w_3}$$

(2.24)

$a_{2p}$ and $a_{3p}$ (Kg/mol) are the parameters directly related to the second virial coefficient $A_{2p}$ and $A_{3p}$ (mL·mol / g²):
and $M_2$, $M_3$, $M_p$ are the molecular weight of polymer 2, polymer 3 and the protein respectively. Superscripts $T$ and $B$ denote top phase and bottom phase respectively. $A_{2p}$ and $A_{3p}$ can be directly obtained by laser light scattering technique as well as membrane osmometry measurements. Substituting Equation 2.24 and 2.25 into 2.23, we have:

$$
\ln K_p = 2A_{2p}M_p\left(\frac{w_2}{1-w_2-w_3}\right)^T - \left(\frac{w_3}{1-w_2-w_3}\right)^T + 2A_{3p}M_p\left(\frac{w_3}{1-w_2-w_3}\right)^T - \left(\frac{w_3}{1-w_2-w_3}\right)^T + \frac{z_pF}{RT} \Delta \varphi
$$

(2.26)

When the polymer concentrations are low, one can have:

$$
\ln K_p = 2A_{2p}M_p(w_2^T) + 2A_{3p}M_p(w_3^T) + \frac{z_pF}{RT} \Delta \varphi
$$

(2.27)

King et al. (1988) reported good predictions for partitioning only by adding the electrostatic term $\frac{z_pF}{RT} \Delta \varphi$ in Equation 2.23. Equation 2.22 and 2.27 are more alike than they first appear (Walter et al., 1991). Equation 2.23 has neglected the higher order terms in the virial expansion. For parallel tie-lines, often a reasonable approximation, the polymer 3 concentration difference can be considered proportional to the polymer 2 concentration difference in Equation 2.27. Therefore, if we neglect the second order term in Equation 2.22, which is often reasonable when the concentration difference is small, Equations 2.22 and 2.27 both approximate to:

$$
2A_{2p} = \frac{a_{2p}}{M_2 M_p}
$$

(2.25)

$$
2A_{3p} = \frac{a_{3p}}{M_3 M_p}
$$
\[
\ln K_p = a(w_2^b - w_2^f) + \frac{z_p F}{RT} \Delta \phi
\]  

(2.28)

where \(a\), a combination of the second virial coefficients \(A_{2p}\) and \(A_{3p}\), indicates the magnitude of the second virial coefficients between the protein and the two polymers in the phase system:

\[
a = \frac{2M_p}{(1 - w_2 - w_3)\omega_{\text{mix}}[A_{2p} + A_{3p} \left(\frac{w_3^b - w_3^f}{w_2^b - w_2^f}\right)]} 
\]  

(2.29)

Other models have tried to incorporate \(K_p\) more strictly based on different molecular interactions (Haynes et al., 1989, 1993; Peng et al., 1995; Gaube et al., 1993). Haynes et al. (1993) applied the mean-spherical approximation (MSA) model of integral-equation theory in statistics mechanics to account different interactions in the aqueous-two phase systems. A free energy model under the McMillan-Mayer framework was developed. Five types of interaction were considered and the short-range force between molecules was calculated by subtracting the charge-charge interaction term and hard sphere interaction term from the second virial coefficient, which was measured by laser light scattering. Good predictions of the phase diagram and protein partition coefficient were reported. However, the application of the model is limited because of the difficulties in determining different parameters for each different molecular interaction.
4. The Second Virial Coefficient and Laser Light Scattering

As we have seen, the virial expansion model (King et al., 1988) has become very attractive because of its sound thermodynamic base and its applicability: however, before we can predict the protein partitioning from the model, we have to obtain the second virial coefficients between the protein and polymers, which characterize the molecular interactions between the protein and polymers in the solution. Details of the virial expansion model and McMillan-Mayer solution theory can be found in Hill’s book (1962). The basic idea of McMillan-Mayer theory is to apply virial expansion to a solution by treating the ‘solute in the solvent’ exactly like treating the ‘gas in the vacuum’. The expression of solution osmotic pressure then can be derived as:

$$\frac{\Pi}{kT} = \rho_2 + \sum_{n=2}^{\infty} B_n(\mu_1, T) \rho_2^n$$

(2.30)

where $\Pi$ is the osmotic pressure; $\rho_2$ is the volume concentration of solute 2; $\mu_1$ is the chemical potential of solvent 1 and $B_n$ is the nth virial coefficient.

$B_2$, the second virial coefficient, describes the interactions between two solute molecules in the solution. The $B_2$ term in Equation 2.30 represents the contributions of two-molecule interaction to the osmotic pressure and is often used to characterize the thermodynamic properties of dilute protein solutions. Statistical mechanics gives the expression for $B_2$ as (Hill, 1962):

$$B_2(a^0, T) = -\frac{1}{2} \int_0^\infty \{ \exp[-w(r, a^0, T)/kT] - 1 \} 4\pi r^2 dr$$

(2.31)
where \( a_1^0 \) is the activity of the pure solvent; \( T \) is the temperature; \( v(r, a_1^0, T) \) is the potential of the mean force between two molecules. Hill (1962) noted that, when including the electrostatic potential to calculate the second virial coefficient of a charged solute such as protein, the solvent and all the other ions in the solution have to be considered as ‘solvent media’. The effect of ions is taken account by adding a screening effect without which the integration of the coulombic force potential would diverge.

Vilker et al. (1981) summarized theoretical expressions for five different interaction force potentials and their respective screening factors. The five are: dispersion, charge-charge, charge-dipole, dipole-dipole effect, and volume exclusion. The second virial coefficients of albumin at three pH values were calculated from these expressions and Equation 2.31 and the predicted osmotic pressure was compared to the experiment results. Only qualitative agreement was obtained which indicates that either a three-term virial equation is insufficient when the protein concentration is high, or that the interaction potentials used oversimplified the actual protein intermolecular interactions.

The second virial coefficient can also be measured by laser light scattering. The theory and practice of light scattering is systematically described by Kratochvil (1987). In multi-angle laser light scattering, a laser beam passing through a solution is scattered by the solute and the intensities of the light at different scattering angles are measured. These intensities can be related to the molecular size and virial coefficient using a Zimm plot:

\[
\frac{R_0}{Kc} = M_i P(\theta) - 2 A_0 c_i M_i^2 P^2(\theta)
\]  
(2.32)
where $R_0$ is the Rayleigh ratio, directly related to the scattered light intensities at angle $\theta$. implying the light scattered in excess of the solvent; $c_i$ is the concentration of the solute in g/mL; $M_i$ is the molecular weight of the solute; $P(\theta)$ is a form factor related to the distance between scattering centers; and $A_n$ is the second virial coefficient of the solute in that solution.

By extrapolating to zero angle, $P(\theta)$ goes to 1 and Equation 2.32 can be simplified as:

$$\frac{Kc_i}{R_0} = \frac{1}{M_i} + 2A_n c_i$$

(2.33)

The second virial then can be calculated from the slope of a plot of $\frac{Kc_i}{R_0}$ vs. $c$, while the molecular weight of the solute can be obtained from the intercept.

Haynes et al. (1989, 1991, 1993) utilized the light scattering technique intensively and obtained the second virial coefficients of several proteins and polymers in different buffers. Comparison of the results from light scattering to those from other methods such as membrane osmometry and differential vapor pressure have been reported and it is believed that light scattering provides more accurate results (Haynes et al., 1989, Rathbone et al., 1990). However, overall the data from light scattering are still very limited. All these data were obtained under specific systems with specific conditions, such as temperature, salt type, and ionic strength so they can only be applied to the same systems with the same conditions. Very few results of applying the measured second virial coefficient to determine the
interaction molecular force in the solution have been reported (Coen et al., 1996; Haynes et al., 1990; Coen et al., 1995; Pjura et al., 1995).

Finally, as we have seen the importance of electrostatic effects from Equation 2.12, we would like to start our studies of the protein partitioning in aqueous two-phase systems from the protein charge effects on its partitioning behavior. Our strategy is to obtain the partitioning behavior of the same protein with different surface charge under the same pH. Therefore, we produced two series of genetically-engineered charge modifications of bacteriophage T4 lysozyme expressed in E. coli. The two series, one in the form of charged-fusion tails and the other in the form of charge-change point mutations, provided matching net charges but very different polarity. By examining the partitioning behavior of these T4 lysozyme mutants, we hope to find how different protein surface charge, as well as different charge distribution brought by different genetic engineering methods, will affect the protein partitioning.

On the other hand, the same series of T4 lysozyme would allow us to investigate the other important element of the electrostatic effect, the electrostatic potential difference between the two phases, and the non-electrostatic effects lumped in the $K_q$ term of Equation 2.12. Comparing the partitioning behavior of the T4 lysozyme mutants in phase systems with different polymer concentrations, different salts and salt concentrations under the same pH, we attempt to understand how different phase systems will have different electrostatic potential difference and different non-electrostatic effects.
To understand the non-electrostatic effects, as we have seen in either Flory-Huggins model or virial expansion model, we will also need to investigate the molecular interactions between proteins and polymers. This is why we need to use laser light scattering and McMillan-Mayer theory to measure and calculate the second virial coefficient, which describes the interactions between two solute molecules in the solution. Examination of the second virial coefficient of different T4 lysozyme mutants will help us understand how change of protein surface properties, i.e. surface charge and charge distribution by genetic engineering methods, will affect molecular interactions. The latter is very important in that eventually, it is the molecular interactions that determine the protein separation not only in aqueous two-phase systems, but also in precipitation, membrane separation, and all the other separation processes. Therefore, studies of different molecular interactions caused by changes of protein surface properties through genetic engineering methods will provide us guidelines on how to improve the target protein selectivity by genetic engineering methods not only in aqueous two-phase systems, but also in all the other processes.
CHAPTER 3. CONTRIBUTION OF PROTEIN CHARGE TO PARTITION IN AQUEOUS TWO-PHASE SYSTEMS

A paper submitted to Biotechnology & Bioengineering

Weiyu Fan, Ufuk Bakir and Charles E. Glatz

Department of Chemical Engineering, Iowa State University, Ames, IA 50011

ABSTRACT

Protein partitioning in aqueous two-phase systems based on phase-forming polymers is strongly affected by the net charge of the protein, but a thermodynamic description of the charge effects has been hindered by conflicting results. Many of the difficulties could be because of problems in isolating electrochemical effects from other interactions of phase components.

We explored charge effects on protein partitioning in poly(ethylene glycol)-dextran two-phase systems by using two series of genetically-engineered charge modifications of bacteriophage T4 lysozyme produced in E. coli. The two series, one in the form of charged-fusion tails and the other in the form of charge-change point mutations, provided matching net charges but very different polarity. Partition coefficients of both series were obtained and interfacial potential differences of the phase systems were measured. Multi-angle laser light scattering measurements were also performed to determine second virial coefficients. A semi-empirical model accounting for the roles of both charge and non-charge effects on
protein partitioning behavior is proposed and the results predicted from the model are compared to the results from the experiments.

INTRODUCTION

Since first used for bioseparations by Albertsson (Albertsson, 1956), aqueous two phase systems have drawn interest because they offer mild conditions, short processing times and ease of scale-up (Albertsson, 1986; Walter et al., 1985; Zaslavsky, 1995).

Considerable advances in the theory of aqueous two-phase systems and partitioning phenomena have been made in the last ten years (Abbort et al., 1990; Albertsson, 1986; Baskir et al., 1987, 1989; Diamond and Hsu, 1989, 1990; Eiteman et al., 1994; Forciniti and Hall, 1990; Forciniti et al., 1991, 1992; Franco et al., 1996a, b; Großmann et al., 1995; Hartounian et al., 1994; Haynes et al., 1989, 1991, 1993; King et al., 1988; Peng et al., 1995; Walter et al., 1985, 1991; Zaslavsky, 1995). Interactions determining protein partitioning in phase systems have been characterized (Albertsson, 1986). Among them, electrostatic interaction plays an important role in the protein partitioning in that protein usually carries net charge and it has been found that an interfacial electrostatic potential difference exists between two phases in a phase system (Johansson, 1985, 1974a, b; Albertsson, 1986).

Albertsson (1986) proposed an expression for the partition coefficient when the electrochemical force dominates the partitioning process:
\[ \ln K_p = \ln K_0 + \frac{z_p F}{RT} \Delta \varphi \]  

where \( K_0 \) is the partition coefficient in the absence of interfacial potential difference; \( z_p \) is the protein net surface charge; \( \Delta \varphi \) is the electrostatic potential difference as bottom minus top; \( F \) is Faraday’s constant and \( T \) is the temperature.

Use of this equation for prediction of the partition coefficient for a given protein in a phase system would require a model for \( K_0 \). More often, the potential difference has been calculated from the slope of a plot of measured \( \ln K_p \) versus \( z_p \). Using the latter method, while varying \( z_p \) by changing pH, potential differences in the range of 0-20 mV were found (Forciniti et al., 1992, 1991; Hartounian et al., 1994; Haynes et al., 1989; Johansson, 1974a, 1985; Luther and Glatz, 1994, 1995; Schluck et al., 1995a, b). However, a linear relationship between \( z_p \) and \( \ln K_p \) was not always obtained. Nonlinearity can result from pH dependence of \( \Delta \varphi \), protein conformation and phase compositions. (Forciniti et al., 1992; Schluck et al., 1995a, b; Zaslavsky, 1995). The latter changes would directly alter \( K_0 \).

Previous studies (Luther and Glatz, 1994) have employed genetic engineering techniques to alter a protein net charge so that the role of charge on partitioning could be investigated without modifying the phase system. A series of \( \beta \)-galactosidases (Zhao et al., 1990) with different lengths of polyasparate fused to the C-termini and T4 lysozymes (Dao-Pin et al., 1991) with charge-change point mutations were obtained and their partition coefficients in aqueous two-phase systems were measured. It was found that the partitioning behavior of T4 lysozyme mutants agreed only qualitatively with Equation 1, while for the \( \beta \)-
galactosidase fusions, a linear relationship between \( \ln K_p \) and \( z_p \) was not observed. One explanation is that Equation 1 does not properly account for electrostatic effects. Alternatively, non-electrostatic factors, lumped into the \( K_o \) term in Equation 1, may differ among the mutants and influence the protein partitioning. Particularly when adding long fusion tails to the protein, the size, conformation, hydration and dipole moment could be changed leading to different molecular interactions in the phase solution.

Because of the very different properties of T4 lysozyme and \( \beta \)-galactosidase, it was difficult to compare charge changes effects brought about by point mutations with those from fusions. In this research, we used fusions of polyarginine tails to the C-terminus of the T4 lysozyme triple point mutant, which provided sets of point mutants and fusions of matching net charge. Comparing the partition coefficients of the two versions of lysozymes allows comparison of "charge" and "tail" effects. Light scattering adds independent data on the effects of the same changes on the \( K_o \) term. A semi-empirical model is proposed which accounts for the partitioning behaviors of these proteins in a series of phase systems.

**THERMODYNAMIC FRAMEWORK**

Even though it has been argued (Zaslavsky, 1995) that an interfacial potential difference won’t affect the partitioning of ionic solutes because of the offsetting effects on transfer of the counterion to preserve electroneutrality, the potential difference term in Equation 1 does indicate the contribution of pure electrical potential to the chemical potential of an ionic species in phase solutions. As we contrast those effects appearing in \( K_o \) with
those in the electrostatic term, it is clear that the latter term results from dividing the electrochemical potential into chemical and electrical contributions (Newman, 1991; Haynes et al., 1991):

\[ \tilde{\mu}_i = \mu_i^{\text{chem}} + z_i F \phi \]  

(2)

where \( \tilde{\mu}_i \) is the electrochemical potential of species \( i \); \( \mu_i^{\text{chem}} \) is the chemical potential of species \( i \); \( z_i \) is the charge of the ionic species and \( \phi \) is the "purely" electrostatic potential. Hence, the usual formulation

\[ \Delta \mu_i = \mu_i^0 + RT \ln m_i \gamma_i \]  

(3)

becomes

\[ \tilde{\mu}_i = \mu_i^0 + RT \ln m_i \Gamma_i + z_i F \phi \]  

(4)

where \( \mu_i^0 \) is the chemical potential at the standard state; \( \gamma_i \) is the activity coefficient; \( m_i \) is the concentration, \( z_i \) is the charge of the ionic species, \( \Gamma_i \) is a "chemical" activity coefficient and \( \phi \) is the "purely" electrostatic potential. Equating \( \mu_i \) for each phase, the result is in the form of Equation 1:

\[ \ln K_i = \ln \frac{m_i^T}{m_i^B} = \ln \frac{\Gamma_i^B}{\Gamma_i^T} + \frac{z_i F}{RT} (\varphi^B - \varphi^T) \]  

(5)

so \( K_0 \) is seen to account for chemical effects caused by non-electrostatic interactions and \( \Delta \varphi = \varphi^B - \varphi^T \), is the electrostatic potential difference. Superscripts \( T \) and \( B \) denote top phase and bottom phase respectively.
Different models have been developed for $\mu_{\text{chem}}$ and thus account for the non-electrostatic interactions, i.e. the $\ln \frac{\Gamma_i^b}{\Gamma_i}$ term in Equation 5 (King et al., 1988; Brooks et al., 1985; Baskir et al. 1987, 1989; Haynes et al., 1991, 1993). The Flory-Huggins model (a lattice model) and the osmotic virial expansion model are among the most important. Brooks et al. (1985) first applied the Flory-Huggins model (Flory, 1953, Huggins, 1942) to protein partitioning in aqueous two-phase systems. Diamond and Hsu (1990) compared the magnitude of the five terms in the Flory-Huggins model and reduced the expression to two terms which were combined with the electrostatic term:

$$\ln K_p = k_1(w_2^b - w_2^T) + k_2(w_2^b - w_2^T)^2 + \frac{z_p F}{RT} \Delta \phi$$

(6)

where $k_1$ and $k_2$ contain the Flory-Huggins interaction parameters, the molar volume ratios and the protein molecular weight; $w_2^b$ and $w_2^T$ are weight fractions of polymer 2, i.e. PEG, in the bottom and top phases, respectively. The virial expansion model expression for $\mu_{\text{chem}}$ after truncation beyond the second virial coefficient term (Edmond and Ogston, 1968; Hill, 1962) and after addition of the electrostatic term (King et al., 1988), leads to:

$$\ln K_p = a_{2p}(m_2^a - m_2^T) + a_{3p}(m_3^a - m_3^T) + \frac{z_p F}{RT} \Delta \phi$$

(7)

where $m_2^T$, $m_2^a$ and $m_3^T$, $m_3^a$ are the molal concentrations of polymer 2 and 3 in the top phase and bottom phase; and $a_{2p}$ and $a_{3p}$ are the parameters directly related to the second virial coefficients $A_{2p}$ and $A_{3p}$, in units of mL·mol / g²:
\[ 2A_{2p} = 1000 \frac{a_{2p}}{M_2 M_p} \]

\[ 2A_{3p} = 1000 \frac{a_{3p}}{M_3 M_p} \]

\( M_2, M_3, M_p \) are the molecular weights of polymer 2 and 3 and the protein respectively.

Under the conditions realized in the experiments reported here, Equation 6 and 7 can be reconciled into a simpler expression (Walter et al., 1991). Converting Equation 7 into weight fractions and assuming parallel tie-lines, which makes \( (w_f^g - w_f^T) \) proportional to \( (w_i^g - w_i^T) \), gives

\[ \ln K_p = a(w_i^g - w_i^T) + \frac{F}{RT} \Delta \phi \]  
(9)

where

\[ a = \frac{2M_p}{(1 - w_2 - w_3)} \times \text{Overall} \left[ A_{2p} + A_{3p} \frac{(w_i^g - w_i^T)}{(w_i^g - w_i^T)} \right] \]  
(10)

captures the virial coefficient effects. Where \( \Delta \phi \) is proportional to the tie-line length (TLL) (Diamond et al., 1989; Hartounian et al., 1994; King et al., 1988), defined as

\[ \text{TLL} = \left[ (w_i^g - w_i^T)^2 + (w_i^g - w_i^T)^2 \right]^{0.5} \]  
(11)

and tie-lines are parallel, \( \Delta \phi \) is also proportional to \( (w_i^g - w_i^T) \). Thus, the partition coefficient can be expressed as:

\[ \ln K_p = (w_i^g - w_i^T)(a + b \Delta \phi) \]  
(12)

where

\[ b = \frac{(F / RT) \Delta \phi}{w_i^g - w_i^T} \]  
(13)
The linear relationship between partition coefficient and protein net charge can be rearranged to:

\[ \frac{\ln K_p}{(w^p_w - w^p_f)} = a + b z_p \]  \hspace{1cm} (14)

where \( a \) and \( b \) can be determined from the intercept and slope of a plot of \( \frac{\ln K_p}{(w^p_w - w^p_f)} \) vs. \( z_p \).

Prediction of partitioning behavior in a two phase system via Equation 7 or 14 requires information on \( \Delta \phi \) and the virial coefficients, \( A_{p} \) and \( A_{rp} \). The potential difference can only be approximated by measurements with Ag/AgCl electrodes (King et al., 1988; Haynes et al., 1989, 1991; Luther and Glatz, 1994), as it is not directly measurable (Guggenheim, 1959). The potential difference measured with Ag/AgCl electrodes is actually the difference of the quasi-electrostatic potential \( \Phi \) of a reference ion (Haynes et al., 1991; Newman, 1991) defined by:

\[ \tilde{\mu}_r = \mu^0_r + RT \ln(m_r \gamma_r) = RT \ln m_r + z_r \Phi \]  \hspace{1cm} (15)

where \( \tilde{\mu}_r \) is the electrochemical potential of the reference ion \( r \); \( \mu^0_r \) is the chemical potential at the standard state; \( \gamma_r \) is the activity coefficient and \( m_r \) is the molality of the reference ion.

The required second virial coefficients can also be measured by multi-angle laser light scattering (Haynes et al, 1989, 1991, 1993; King et al, 1988; Kratochvil, 1987; Rathbone et al., 1990). Comparing the value of \( a \) regressed from Equation 14 to the value determined from light scattering will provide insight as to how different molecular interactions influence the protein partitioning behavior in an aqueous two-phase system. More recently, the phase
equilibrium calculations based on virial coefficients have been directly linked to specific intermolecular interactions of the components (Haynes et al., 1993; Gaube et al., 1993; Coen et al., 1995). Such a step requires a larger number of molecular properties not readily available for different phase systems and proteins. We have applied this approach to the $A_n$ of the proteins used in this study but not the $A_g$ required here (Fan and Glatz, 1997).

**EXPERIMENTAL**

*Materials*

Polymers. Poly(ethylene glycol) (PEG-3350, $M_w=3350$) and dextran (Dextran-40, $M_w=39100$) were purchased from Sigma.

Proteins. Two series of genetically engineered T4 lysozymes were produced. One is the mutant series (strains provided by B. W. Matthews, University of Oregon) that has been modified through site-directed mutagenesis to replace lysine residues with glutamic acid (Dao-Pin, 1991). Each mutation results in an expected reduction of 2 units of charge at neutral pH. Using the numbering notation of Dao-Pin to indicate the mutation site, mutants containing one, two and three mutations are denoted as K16E, K16/135E, and K16/135/147E, respectively; “K” and “E” indicate that lysine is replaced by glutamic acid; WT denotes wild-type. The second “fusion” series consists of 2 fusions of polyarginine tails containing 2 and 4 arginines (designated U1 and U2, respectively) to the carboxyl terminus of mutant T4 lysozyme K16/135/147E. The genes for all the T4 lysozymes are carried on the expression vector pHN1403, which has been transformed into *E. coli* strain RR1.
Full enzymatic activity is observed for all the mutants and fusions. The estimation of charge of T4 lysozyme wild type and mutants (Table 1) has been reported previously (Luther and Glatz, 1994). U1 has the same charge as that of double mutant K16/135E while U2 has the same charge as that of the single mutant K16E. Figure 1 shows that native gel electrophoresis confirms the altered charge and the equivalencies of the corresponding mutants. Modeling of the configuration of the U1 fusion protein starting from the wild type T4 lysozyme structure using the Swiss-Model (an automated protein modeling server running at the Geneva Biomedical Research Institute, Glaxo Wellcome Research and Development, S.A., Switzerland) predicts that the fusion tail protrudes from the protein surface. The same conclusion on tail configuration was obtained for similar fusion tail proteins using size exclusion chromatography (Niederauer et al., 1993).

Salts. All the salts were ACS Certified and obtained from Fisher Scientific. (Fair Lawn, NJ)

Methods

Fusion protein construction

The cassette mutagenesis technique was used for the construction of the polyarginine fused, T4-lysozyme proteins. The triple mutant of the T4-lysozyme gene was in the plasmid pHN1403 between BamHI and HindIII restriction enzyme sites (Muchmore et al., 1989; Dao-Pin et al., 1991; Potee et al., 1991). The gene was cleaved from the plasmid by
HindIII/BamHI double digestion, cloned into pBluescript SKII(+) to facilitate the genetic protocols and then transformed into *E. coli* XL1-blue.

To add the polyarginine tails to the C-terminal end of the protein, two unique restriction enzyme sites were searched enclosing the C-terminal. A HindIII site, immediately downstream of the C-terminal, and a *MluI* site, 12 bp upstream of the C-terminal, were found. Two complementary single-stranded oligonucleotides containing the removed C-terminal portion and a polyarginine tail containing either two or four arginines were prepared. Also, to facilitate the identification of the tail after cloning, *PstI* and *SmaI* restriction enzyme sites were included after the stop codon to two and four arginine containing tails, respectively. By annealing these complementary oligonucleotides, new cassettes were formed. After removing the original cassette by double digestion with HindIII and *MluI*, new cassettes were cloned back into the T4-lysozyme gene in pBluescriptII one by one. The generation of the fusion proteins was verified by checking the presence of the unique restriction enzyme sites for each tail after the stop codon and also by gene sequencing.

To express the fusion protein, the corresponding genes were isolated by double digestion with *BamHI* and HindIII. After removing the original triple mutant T4-lysozyme gene from pHN1403, the fusion protein genes were cloned into pHN1403, the presence of the tails were again confirmed by restriction enzyme digestion and transformed into a *E. coli* strain, RR1.

Protein production and activity assay.
T4 lysozyme production was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) and the enzymes were purified as described by Luther and Glatz (1994). Lysozyme activity was measured by following the clearing of a *Micrococcus lysodeikticus* cell suspension versus time (Parry et al., 1965). Protein concentration for the lysozyme mutants and fusions were determined by measuring the absorbance at 280 nm, using an extinction coefficient of 1.28 (Dao-Pin, et al., 1991).

**Aqueous two-phase systems**

Four phase systems, System A, B, C and D were prepared from stock solutions of dextran, polyethylene glycol, mono- and dibasic potassium phosphate and deionized water at the polymer concentrations shown in Figure 2 (PEG concentrations of 8.5, 9.5, 10.5 and 11.5% w/w and dextran concentration 8, 9, 10, 11% w/w respectively) in 0.01M potassium phosphate buffer at pH 7.25. The phase systems were mixed by vortexing for 10 seconds and equilibrated at room temperature (23°C) for 2 hours. In order to speed up the phase separation, phase systems were centrifuged at 3000 g for 5 minutes if after two hours the phase separation was still not complete. PEG and dextran concentrations in the top and bottom phases were measured by size exclusion chromatography (TosoHaas TSK-GEL G3000PW). The samples were diluted to a concentration level appropriate for the capacity of the column.

**Protein partition coefficient measurements**
Stock protein solutions (ca. 1 mg/mL) were prepared by dialyzing the protein solutions against 0.01M potassium phosphate buffer (pH 7.25). For all four phase systems, bulk PEG and dextran solutions with different calculated amount were aliquoted into polystyrene tubes and 1 g protein solution was added. Water was added to the final weight of 8 g. A blank phase system was also prepared by adding 1 g of 0.01 M potassium phosphate buffer in place of the 1 g protein solution. The systems were mixed by vortexing for 10 seconds, allowed to separate for two hours, and then were centrifuged at 3000 g for 5 minutes. In no case was any precipitation observed at the interface. Top phase samples were collected by a pipette while those from the bottom phase were obtained by piercing the bottom of the tube and allowing to drain. The protein concentration of each phase was determined. The mass balance of added protein was checked and showed no protein loss during separation. All partitioning experiments were performed in triplicate and at room temperature, 23°C. A preliminary comparison of partitioning behavior with protein concentration 2x smaller and larger than this showed agreement within 6%.

Potential measurements

Electrodes made of a 1.5 mm glass capillary filled with 2% agar in 3 M KCl were used to measure $\Delta\Phi$ as described by Luther and Glatz (1994). At least three measurements of each phase system were made. All potential differences reported are expressed as bottom phase minus top phase.
Multi-angle laser-light scattering

Light scattering measurement were made using the DAWN-B light scattering instrument (Wyatt Technologies Corporation, Santa Barbara, CA), a batch system employing a laser light beam at 632.8 nm and measuring the intensity of light scattered at 15 different angles ranging from 23° to 128°. Refractive index increments were measured on an OPTILAB 903 Differential Refractometer (Wyatt Instruments, Inc.) also operating at 632.8 nm. All the polymer sample buffers were prepared using deionized water (Barnstead, NanopureII water treatment system, Barnstead CO., Boston, MA) and then filtered (0.2 μm, Gelman Sciences, Ann Arbor, MI). Protein samples were diluted and dialyzed overnight against filtered (0.02 μm Anotop syringe tip filters, Whatman) potassium phosphate buffer (0.01 M, pH 7.25). Glass scintillation vials were used as sample cells in the DAWN-B instrument. All the measurements were performed at room temperature (23°C).

A graphical technique, the Zimm plot, was employed in order to extrapolate to the values of the zero angle Rayleigh ratio at different concentration levels of single polymer or protein solutions (Kratochvil, 1987). The second virial coefficients, $A_i$, and molecular weight were obtained directly from the slope and intercept of the plot at zero angle:

$$\frac{K_c}{R_0} = \frac{1}{M_i} + 2A_i c_i$$

where $R_0$ is the Rayleigh ratio, directly related to the scattered light intensity at angle $\theta$; $c_i$ is the concentration of the solute in g/mL; $M_i$ is the molecular weight of the solute; and $K'$ is the single-component optical constant.
The polymer-protein second virial cross coefficient, $A_y$, can be determined from the Rayleigh ratio of a solution containing a protein and a polymer:

$$K'' \frac{(c_i + c_j)}{R_0} = h(c_i + c_j) + l$$

(17)

where $c_i$ and $c_j$ are concentrations of the protein and polymer in g/mL in the solution: $K''$ is the multi-component optical constant; $l = \frac{1}{v_i^2 M_{w_i} w_i + v_j^2 M_{w_j} w_j}$; and

$$h = \frac{2(v_i^2 M_{w_i} w_i^2 A_i + 2v_i v_j M_{w_i} M_{w_j} A_j + v_j^2 M_{w_j} w_j^2 A_j)}{(v_i^2 M_{w_i} w_i + v_j^2 M_{w_j} w_j)^2}.$$

RESULTS AND DISCUSSION

Figure 2 shows the phase compositions of four PEG 3350 and Dextran-40 phase systems. It also shows that "parallel tie-lines" is a good approximation. System A with the lowest polymer concentration has the smallest concentration difference between top and bottom phase. TLL of the four systems are also listed with Figure 2.

The measured interfacial potential differences (Table 2) show the same linear trend with PEG concentration difference (or TLL) (Figure 3) as reported by others (King et al., 1988; Haynes et al., 1989, 1991; Luther and Glatz, 1994). The potential difference was the expected zero with the two electrodes immersed in the same phase. Electrode junction potential errors are less than 10% of the measured electrostatic potential difference, as
estimated from the Henderson formula (Haynes et al., 1991). Furthermore, the two such junction potentials would tend to cancel each other. The addition of protein should not affect the potential difference, as in our experiments the protein concentration (typically 0.005 mM) as well as the concentration of charges provided by the protein (≤ 0.05 mM) is much lower than the salt concentration (10 mM) (Johansson, 1974a).

However, \( \Delta \Phi \), the potential difference measured with Ag/AgCl electrodes, is the quasi-electrostatic potential difference which differs from the pure electrical potential difference \( \Delta \phi \) in Equation 1 and 9. This may account for the inconsistency in using \( \Delta \Phi \) in Equation 1 to describe partitioning behavior. Figure 4 shows the partition coefficients of the series of T4 lysozyme point mutants in the four aqueous two-phase systems; these data are replotted in the form of Equation 1 in Figure 5 and show the expected form. As the interfacial potential difference increases or the protein net charge increases, the partition coefficient decreases. However, using Equation 1 to calculate the interfacial potential difference from the slopes in Figure 5 gives values of interfacial potential difference different from those measured (Table 2).

One can not apply Equation 1 to different phase systems because Figure 5 clearly indicates that the four different phase systems have four different \( K_0 \) values. Equation 9 (or 12), on the other hand, takes into account the non-electrostatic effects on the protein partition behavior caused by the different proteins and different polymer concentrations in the phase systems. Figure 6 shows the linear relationship obtained for all four T4 lysozyme mutants in the four phase systems when plotted according to Equation 14. Furthermore, for the same
protein, the value of \( \frac{\ln K_p}{(w_2^p - w_2^T)} \) falls in the same range for all four phase systems, indicating that the model accounts for the effect of polymer concentration. Linear regression of the partitioning data for the four proteins in the four PEG3350/dextran-40 systems gives:

\[ \ln K_p = (w_2^p - w_2^T)(5.9 + 1.4z_p) \]  \hspace{1cm} (18)

where the slope \( b \) can be used to calculate the relationship between the interfacial phase potential difference and the polymer concentration difference, \( \Delta w \), from Equation 13. Again, the calculated results yield a potential difference value different from the value measured from the experiment (Table 2). Equivalently, the prediction of protein partition coefficient from Equation 9 using \( a = 5.9 \) and the measured potential difference is not very satisfactory (Figure 5).

Equation 18 can not be used to predict the partitioning behavior of the T4 lysozyme fusion series in the four phase systems. Figure 7 shows that the fusion tail lysozymes, U1 and U2, have different partition coefficients than the corresponding double and single point mutants even though U1 has the same net charge as that of double mutant and U2 has the same net charge as that of single mutant. Although a linear relationship in the form of Equation 14 can still be observed, Figure 8 shows the values of intercept and slope obtained from fusions in four different phase systems are different from those obtained from point mutants. For T4 lysozyme fusions, we have \( a = 9.4 \) and \( b = 0.37 \). The two different slopes and intercepts each have different implications.
First, the phase potential difference calculated from the slope of Figure 8 (\( b \) value of the fusions) is very different from the potential difference calculated from Figure 6 (\( b \) value of the point mutations) (Table 2), which indicates the electrostatic term \( \frac{zF}{RT} \Delta \phi \) in the equation cannot be used to describe the electrostatic effect brought by the charged fusion tails. It has been pointed out (Luther and Glatz, 1994; Albertsson, 1986) that this electrostatic term strictly applies only to a point charge, though it appears to be a good approximation when the charge is distributed evenly over the surface. The latter seems reasonable for the point mutants, so that the protein is directly affected by the bulk phase potential and the double layer effects can be ignored (Neogi, 1993); however, this would not apply for the high charge density of the fusion tails. On the other hand, it also might have implied that the addition of different lengths of charged fusion tail not only has changed protein net surface charge, but also has changed protein-polymer molecular interactions in the phase solution, which will cause different protein-polymer second virial coefficients and \( K_\theta \) for each different fusion protein even in the same phase system.

Second, the difference in the intercepts, \( a \), implies that the interactions between fusions and polymers in the solution are different from those between mutants and polymers since \( a \) reflects the magnitude of the second virial coefficients between the two polymers and the protein, depicting the interaction forces in the phase solution (Equation 10). Fusing a charged tail to the protein molecule obviously will change the protein dipole moment which brings the need to consider the dipole-dipole and dipole-charge interactions between the polymer and protein molecules in the phase solution. A high charge-density tail at the end of
the protein molecule will affect the hydration, which could in turn affect its interaction with the polymer molecules through competition for water molecules.

Therefore, even though the point mutations and charged tail fusions can bring the protein to the same net charge, very different protein-protein and protein-polymer molecular interactions result, leading to different partitioning behavior.

The second virial coefficients of PEG, dextran, and six T4 lysozymes in 0.01 M K-PO$_4$ at pH=7.25 obtained from multi-angle laser light scattering are reported in Table 3. The second virial cross coefficients, $A_{ij}$, between each of the six lysozymes and each of the two polymers are shown in Table 4. The differences in charge and charge distribution among the T4 lysozymes changes both protein-protein and protein-polymer interactions. Positive values of the second virial coefficient indicate that the average force between the two molecules is repulsive, while negative values indicate an attractive mean force between the two molecules (Coen et al., 1995, 1996; Forciniti and Hall, 1990; Haynes et al., 1989, 1991; Pjura et al., 1995). The expectation discussed above, that the addition of a charged tail has increased the dipole-dipole and charge-dipole attractions, is seen in the lower values of $A_{ij}$ between the fusions and the two polymers compared to those between the corresponding point mutants and the same two polymers (Coen et al., 1995, 1996).

Use of these $A_{ij}$ with Equation 10 results in larger calculated partition coefficients for the fusions than for the corresponding point mutation mutants. The same trend is evident from the experiment (Figure 7). However, the predicted magnitudes are not in agreement. The value of $a$ calculated from the $A_{ij}$ and Equation 10 is much larger than the value
regressed from partitioning results in Figure 6 (Table 5). Hence, the magnitude of the concentration difference term in Equations 8 and 9 is much larger than that of the potential difference term in the equations, which results in much smaller partition coefficients than those from experiments. One reason could be the difficulties in determining the $A_{ij}$ because $A_{ij}$ can only be calculated from Equation 14 based on the values of $A_{ii}$, $A_{jj}$ and $h$ measured from light scattering experiments. We have calculated the second virial coefficients of T4 lysozyme point mutations and fusion mutants from McMillan-Mayer theory (Fan and Glatz, 1997). Calculated results agreed with the second virial coefficients measured by laser light scattering at high ionic strengths but not at low ionic strengths. The larger second virial coefficients obtained from light scattering in our experiments imply that because of the large Debye length at low ionic strength, it would be necessary to include higher order virial coefficient as multi-body interactions become more important. Finally, the simplification of treating salt-water as a single-component solvent has been questioned (Forciniti and Hall, 1990). This simplification also prevents the application of the model to different phase systems where different salt concentrations would result in different values of the second virial coefficient.

**CONCLUSIONS**

T4 lysozyme point mutants and fusions of equivalent charge display different partitioning behavior. A model based on a thermodynamic framework and experimental
regression accounted for both non-electrostatic effects caused by the polymer concentrations and electrostatic effects caused by the interfacial potential difference. The model was unable to reconcile the fusion tail behavior where the addition of a charged tail has not only changed the net charge of the protein, but also, most likely, the dipole moment, the electrostatic potential distribution and the hydration. The potential difference regressed from protein partition coefficient data was different from the measured potential difference, as might be expected given the different thermodynamic meanings of the two potential differences. The results from light scattering provided evidence that T4 lysozyme fusions have different protein-polymer interactions in the phase solution from those of mutants. However, applying the measured second virial coefficient from light scattering to the model did not give a good prediction of the protein partition coefficient.

ACKNOWLEDGMENT

We thank Dr. Brian Matthews for providing the cultures to produce the T4 lysozyme and mutants with charge-change point mutations and Dr. Clark Ford of Iowa State University for guidance and facilities in preparing the genetic constructions. This material is based upon work supported by the National Science Foundation under Grant No. BCS-9108583. The government has certain rights in this material. Weiyu Fan was partially supported by the Biotechnology Byproducts Consortium under a contract from the U.S. Department of Agriculture.
Nomenclature

\( a \) = regression parameter in the partitioning model

\( A_{ii} \) = osmotic second virial coefficient of single solute, \( \text{mL mol/g}^2 \)

\( A_{ij} \) = osmotic second virial cross coefficient between two solute molecules \( i \) and \( j \), \( \text{mL mol/g}^2 \)

\( a_j \) = activity coefficient of species \( j \) in a solution

\( b \) = regression parameter in partitioning model

\( c_i \) = concentration of species \( i \) in the solution

\( F \) = Faraday constant

\( h_l \) = regression parameters in the Zimm plot for solutions containing two solutes

\( k_{1,2} \) = parameters in an expression of partition coefficient

\( K \) = partition coefficient

\( K' \) = single-component optical constant in light scattering

\( K'' \) = multicomponent optical constant in light scattering

\( m_i^B \) = molality of species \( i \) in the bottom phase of the phase system

\( m_i^T \) = molality of species \( i \) in the top phase of the phase system

\( M_i \) = molecular weight of species \( i \)

\( N_A \) = Avogadro’s number

\( R \) = gas law constant

\( R_0 \) = Rayleigh ratio

\( T \) = absolute temperature
\[ u_i = \text{mobility of species } i, \text{ determined from Nernst-Einstein equation} \]

\[ w_i^{B} = \text{weight fraction of species } i \text{ in the bottom phase of the phase system} \]

\[ w_i^{T} = \text{weight fraction of species } i \text{ in the top phase of the phase system} \]

\[ x_s = \text{salt concentration} \]

\[ z_i = \text{charge number of species } i \]

**Greek Letters**

\[ \Delta \varphi = \text{electrostatic potential difference between the two phases in a phase system} \]

\[ \Delta \Phi = \text{quasi-electrostatic potential difference, or measured potential difference between the two phases in a phase system} \]

\[ \mu^{\text{chem}} = \text{chemical potential} \]

\[ \tilde{\mu} = \text{electrochemical potential} \]

\[ \nu = \text{refractive-index increment} \]

\[ \theta = \text{light scattering angle} \]

**REFERENCES**


Suggest that Long-Range Electrostatic Interactions Contribute Little to Protein Stability.


Table I. Net charge of T4 lysozyme mutants and fusions at pH=7.25*

<table>
<thead>
<tr>
<th>T4 Lysozyme</th>
<th>Wild</th>
<th>Single</th>
<th>Double</th>
<th>Triple</th>
<th>U1</th>
<th>U2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Mutant</td>
<td>Mutant</td>
<td>Mutant</td>
<td>Fusion</td>
<td>Fusion</td>
<td></td>
</tr>
<tr>
<td>Net charge</td>
<td>9.84</td>
<td>7.84</td>
<td>5.84</td>
<td>3.84</td>
<td>5.84</td>
<td>7.84</td>
</tr>
</tbody>
</table>

*Charge estimated by using a combination of titration, isoelectric focusing, and calculation from the Henderson-Hasselbalch relationship and pK values for amino acids in proteins from Lehninger (1982) (Luther and Glatz, 1994).
Table II. Comparing the phase potential differences (bottom phase-top phase) measured from experiments to those regressed from the models

<table>
<thead>
<tr>
<th>System</th>
<th>System A</th>
<th>System B</th>
<th>System C</th>
<th>System D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \Phi^\ast$ (measured, mV)</td>
<td>-1.82±0.14</td>
<td>-2.47±0.18</td>
<td>-3.51±0.10</td>
<td>-3.80±0.16</td>
</tr>
<tr>
<td>$\Delta \phi$ (from Equation 1 and Figure 5, using partitioning results of point mutations, mV)</td>
<td>-2.6</td>
<td>-3.7</td>
<td>-4.4</td>
<td>-4.9</td>
</tr>
<tr>
<td>$\Delta \phi$ (from Equation 13 and Figure 6, using partitioning results of point mutations, mV)</td>
<td>-2.3</td>
<td>-3.3</td>
<td>-4.0</td>
<td>-4.6</td>
</tr>
<tr>
<td>$\Delta \phi$ (from Equation 13 and Figure 8, using partitioning results of fusions, mV)</td>
<td>-0.61</td>
<td>-0.87</td>
<td>-1.1</td>
<td>-1.3</td>
</tr>
</tbody>
</table>

*± standard deviation of the mean
Table III. Molecular weight and second virial coefficient of single solute, $A_{ii}$, from multi-angle laser light scattering

<table>
<thead>
<tr>
<th>Sample</th>
<th>Published Molecular Weight</th>
<th>Molecular Weight from Light Scattering</th>
<th>Second Virial Coefficient $A_{ii}$ * mol·mL/g²</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEG3350</td>
<td>3350</td>
<td>3500</td>
<td>0.0044</td>
</tr>
<tr>
<td>Dextran-40</td>
<td>39100</td>
<td>38900</td>
<td>0.00069</td>
</tr>
<tr>
<td>Wild Type</td>
<td>18600</td>
<td>18900</td>
<td>0.058</td>
</tr>
<tr>
<td>Single Mutant</td>
<td>18600</td>
<td>18400</td>
<td>0.049</td>
</tr>
<tr>
<td>Double Mutant</td>
<td>18600</td>
<td>17500</td>
<td>0.017</td>
</tr>
<tr>
<td>Triple Mutant</td>
<td>18600</td>
<td>18100</td>
<td>0.015</td>
</tr>
<tr>
<td>U1 Fusion</td>
<td>18800</td>
<td>18500</td>
<td>0.045</td>
</tr>
<tr>
<td>U2 Fusion</td>
<td>19000</td>
<td>19000</td>
<td>0.040</td>
</tr>
</tbody>
</table>

* The $A_{ii}$ value obtained from the curve fitting of the Zimm plot for each individual measurement has standard deviation less than 25%. Each of the above values is an average of three replicates.
Table IV. second virial cross coefficient between a protein and a polymer $A_{ij}$ from multi-angle laser light scattering

<table>
<thead>
<tr>
<th></th>
<th>Second Virial Cross Coefficient $A_{ij}$, Protein - PEG 3350, mol-mL/g^2</th>
<th>Second Virial Cross Coefficient $A_{ij}$, Protein - Dextran-40, mol-mL/g^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>0.031*</td>
<td>0.0016</td>
</tr>
<tr>
<td>Single Mutant</td>
<td>0.023</td>
<td>0.0012</td>
</tr>
<tr>
<td>Double Mutant</td>
<td>0.014</td>
<td>0.0016*</td>
</tr>
<tr>
<td>Triple Mutant</td>
<td>0.010</td>
<td>0.00052</td>
</tr>
<tr>
<td>U1 Fusion</td>
<td>-0.025</td>
<td>-0.0015</td>
</tr>
<tr>
<td>U2 Fusion</td>
<td>-0.036+</td>
<td>0.0035</td>
</tr>
</tbody>
</table>

* value with % standard deviation between 50% and 75%.

+ value with % standard deviation more than 75%.
Table V. Comparison of the value of "a" calculated from the second cross virial coefficients, $A_{ij}$, in table 5 to the value regressed from protein partition coefficient

<table>
<thead>
<tr>
<th>T4 Lysozyme</th>
<th>a Regressed from Partition Coefficient</th>
<th>a Calculated from $A_{ij}$ in Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>5.9</td>
<td>990</td>
</tr>
<tr>
<td>Single Mutant</td>
<td>5.9</td>
<td>750</td>
</tr>
<tr>
<td>Double Mutant</td>
<td>5.9</td>
<td>380</td>
</tr>
<tr>
<td>Triple Mutant</td>
<td>5.9</td>
<td>320</td>
</tr>
<tr>
<td>U1</td>
<td>9.4</td>
<td>-790</td>
</tr>
<tr>
<td>U2</td>
<td>9.4</td>
<td>-</td>
</tr>
</tbody>
</table>

*The a value of U2 not calculated from $A_{ij}$ because variability of replicates exceeded 75%
FIGURE LEGENDS

Figure 1. Native gel of T4 lysozyme point mutants and fusions, pH 4.0. Tri, Dou and Sng denote mutants containing one, two and three point mutations, respectively. Wild denotes wild-type T4 lysozyme and U1 and U2 denote the fusion mutants.

Figure 2. Binodal and tie-lines for PEG 3350/dextran-40 phase systems. Points • represent overall polymer concentrations of A, B, C, D four systems and points ◊ indicate the phase compositions for the four systems used in this work; tie-line length (TLL) of the A, B, C, D systems are 18.5, 24.8, 29.3 and 32.7% (w/w).

Figure 3. Measured potential difference (bottom-top) vs. PEG concentration difference between the two phases.

Figure 4. Partition coefficients of T4 lysozyme mutants vs. tie-line length in the four different phase systems.

Figure 5. Partition coefficient (as \( \ln K \)) of T4 lysozyme mutants, predicted protein partition coefficient from Equation 1 (solid line), and predicted protein partition coefficient from Equation 12 using measured potential difference (dashed line) vs. protein net charge for each of the four phase systems.

Figure 6. \( \frac{\ln K_p}{(w_2^A - w_2^F)} \) vs. protein net charge for all the T4 lysozyme mutants in the four phase systems.

Figure 7. Comparison of partitioning coefficients of T4 lysozyme mutants to those of the fusions of matching charge in the four phase systems.
Figure 8. $\frac{\ln K_p}{(w_2 - w_2^*)}$ vs. protein net charge for the T4 lysozyme triple mutant and fusions in the four phase systems.
(Figure 3)
(Figure 4)
(Figure 5)
(Figure 6)
(Figure 7)
System A
• System B
System C
System D

Protein Net Charge

In K/PEG Concentration Difference

(Figure 8)
CHAPTER 4. PROTEIN PARTITION IN AQUEOUS PEG-DEXTRAN TWO-PHASE SYSTEMS: EFFECTS OF DIFFERENT IONS AND IONIC STRENGTHS

A paper to be submitted to Biotechnology & Bioengineering

Weiyu Fan and Charles E. Glatz

Department of Chemical Engineering, Iowa State University, Ames, IA 50011

ABSTRACT

This work continues our examination of contribution of protein charge to partitioning in PEG-dextran aqueous two-phase systems. The partitioning of T4 lysozyme and its charge-change mutants, both point mutations and fusion tails, in aqueous two-phase systems with NaCl, KCl, K$_2$SO$_4$ and potassium phosphate four salts and different salt concentrations are reported. Both salts and proteins have different distributions between the two phases in the systems with the same overall polymer concentration, but different salts and salt concentrations. T4 lysozyme point mutants and fusion-tail mutants of the same charge display different partitioning behavior. The point mutation mutants were best suited for distinguishing non-electrostatic from electrostatic effects of the different salts on protein partitioning.
INTRODUCTION

Aqueous two-phase systems are typically composed of solutions of two incompatible polymers, such as polyethylene glycol (PEG) and dextran, and provide mild conditions, short processing times, and ease of scale-up that are desirable for purification of biological materials. Abbott et al. (1990) and Walter et al. (1991) have reviewed the advances in the theories and their application of aqueous two-phase systems.

Previously (Fan et al., 1997), we summarized the literature which has sought to isolate the role of protein charge (Albertsson, 1986; Haynes et al., 1989; King et al., 1988; Diamond et al., 1990); our own approach was to use a series of T4 lysozymes mutants differing in net charge and/or charge distribution. We examined the partitioning behavior of charge-change point mutations and charged-tail fusions in PEG-dextran aqueous two-phase systems of differing polymer concentrations. Protein net charge strongly affected protein distribution between the phases, but the difference between point mutations and fusions of the same net charge showed that charge distribution was also a factor. The dependence of the protein partition coefficient, $K_p$, on phase composition and potential difference was described as:

$$\ln K_p = a (w_2^b - w_2^d) + \frac{z_p F}{RT} \Delta \phi$$

where $(w_2^b - w_2^d)$ is the measure of the polymer concentration difference between the two phases; $a$ is an interaction parameter related to the second virial coefficient between the protein and polymers in a phase solution; $\frac{z_p F}{RT} \Delta \phi$ is the electrostatic term as a product of
protein surface charge $z_p$ and $\Delta \varphi$, the electrostatic potential difference between the two phases. The first term on the right side of the equation is derived from the virial expansion model (King et al., 1988) and it is noted that the same form can also be approximated from the Flory-Huggins model (Diamond et al., 1990). The second term is from the classic model developed by Albertsson (1986) for the electrochemical partitioning of proteins in phase systems:

$$\ln K = \ln K_0 + \frac{z_p F}{RT} \Delta \varphi$$  \hspace{1cm} (2)$$

where $\ln K_0$ includes all the other factors affecting the protein partitioning. However, it has been pointed out (Haynes et al., 1991; Fan et al., 1997) that $\Delta \varphi$ in these equations should be the pure electrostatic potential difference, which differs from the potential difference measured by two Ag/AgCl electrodes. The latter has been termed the quasi-electrostatic potential difference (Haynes et al., 1991; Newman, 1991) and has often been taken as an empirical approximation to the former. The shortcomings of this approximation may have accounted for the earlier observed discrepancy between the potential difference regressed from the protein partitioning results using Equation 2 and the potential difference from direct measurement (Fan et al., 1997).

An examination of charge effects on protein partitioning can not be completed without considering salt effects. Different salts and salt concentrations in a phase solution will cause different solvent structures and molecular interactions (Zaslavsky, 1995) and, in turn, change phase separation and protein partitioning behavior.
Past studies of salt effects on protein partitioning (Johansson, 1970, 1974a, b. 1985; Albertsson, 1986; Abbott et al., 1990) have focused on changes in the potential difference, i.e. $\Delta \phi$ in Equation 2, to account for observed changes in partitioning. At low salt concentrations (0.01-0.1 M), the change in protein partitioning with increasing salt concentration was attributed increasing predominance of the added salt in establishing the interfacial potential difference (Johansson, 1970, 1974b; Albertsson, 1986). The alternative, i.e. that the presence of salt affects $K_q$ (Equation 2) has been neglected. Indeed, independence of $K_q$ among systems with different salts and salt concentrations was assumed (Johansson, 1974a, b; Albertsson, 1986) in order to determine the protein isoelectric point from the cross point of curves of protein partition coefficient versus pH in systems with different salts ("cross partitioning").

However, one would expect that different salts and salt concentrations would effect $K_q$ both through direct changes in protein-polymer interactions and indirectly through alteration of polymer-polymer interactions leading to altered phase compositions. Increasing the concentration of water-structure-making ions, such as $SO_4^{2-}$ and $PO_4^{3-}$, will shift the binodal to lower polymer concentrations, and increasing the concentration of water-structure-breaking ions, such as $Cl^-$, will shift the binodal to higher polymer concentrations (Zaslavsky et al., 1988; Huddleston and Lyddiatt, 1990). This may be why it has been very difficult to determine the protein isoelectric point by cross partitioning (Zaslavsky, 1995).

An additional complication in past studies focusing on protein charge and salt effects is that charge is usually altered by adjusting the pH of the phase systems through alteration of
buffer (typically 0.01M sodium or potassium phosphate salt) composition. This brings two major limitations: 1) the protein conformation can be changed with the change of pH (Forciniti et al., 1991; Schluck et al., 1995a, b), and 2) different buffer salts can have significant effects on solvent properties and phase compositions (Forciniti et al., 1991; Zaslavsky, 1995). As a result, the linear relationship between InK and the protein charge as predicted by Equation 2 is not observed (Forciniti et al., 1991; Schluck et al., 1995a, b); therefore, Equation 2 can not be used to separate the influences of salts on K₀ and Δφ.

In this work, we try to separate the effects of protein charge from pH by use of two series of charge-change mutants and to isolate the effects of added salt by using a range of salt concentrations coupled with monitoring of phase composition. Thus, we are able to investigate both electrostatic and non-electrostatic effects caused by different salts and salt concentrations under the same pH value. The aim is to combine manipulation of protein charge and salt content of the phase system to enhance extraction selectivity.

MATERIALS AND METHODS

Proteins

Two series of genetically engineered T4 lysozymes were produced and purified as described by Luther and Glatz (1994). One is the mutant series (strains provided by B. W. Matthews, University of Oregon) that has been modified through site-directed mutagenesis to replace lysine (K) residues with glutamic acid (E) (Dao-Pin et al., 1991). Each mutation results in an expected reduction of 2 units of charge at neutral pH. Using the numbering
notation of Dao-Pin to indicate the mutation site. Mutants containing one, two, and three
mutations are denoted as K16E, K16/135E, and K16/135/147E, respectively; WT denotes
wild-type. The “fusion” series consists of fusions of polyarginine tails containing 2 and 4
arginines (designated U1 and U2, respectively) to the carboxyl terminus of mutant T4
lysozyme 16/135/147E. The genes for all the T4 lysozymes are carried on the expression
vector pHN1403, which has been transformed into *E. coli* strain RR1.

Lysozyme activity was measured by following the clearing of a *Micrococcus
llysodeiktikus* cell suspension versus time (Parry et al., 1965). Full enzymatic activity is
observed for all the mutants and fusions. Lysozyme concentration was determined by
measuring the absorbance at 280 nm, using an extinction coefficient of 1.28 (Dao-Pin et al.,
1991). The estimation of charge of T4 lysozyme wild type and mutants has been reported
previously (Luther and Glatz, 1994). U1 has the same charge as that of double mutant
K16/135E while U2 has the same charge as that of the single mutant K16E. Table 1 lists the
estimated charge numbers of all 6 lysozymes at pH 7.25.

*Aqueous two-phase systems*

Stock salt (ACS Certified, Fisher Scientific, Fair Lawn, NJ) solutions, 2M sodium
chloride and potassium chloride, 0.7 M potassium sulfate, 1M potassium phosphate at pH
7.25 (with $K_H^2PO_4 : K_2HPO_4 = 0.3 : 0.7$). The amount of stock salt solution needed was
calculated according to the final bulk salt concentration of the phase systems. Fresh
poly(ethylene glycol) (PEG-3350, $M_w=3350$) and dextran (Dextran-40, $M_w=39100$)

(purchased from Sigma Chemical Co., St. Louis, MO) stock solutions were prepared and the polymer concentrations determined from the refractive index (Bausch & Lomb refractometer, Rochester, NY). Bulk phase systems with 8.5\% w/w PEG, 8\% w/w dextran, 0.01 M potassium phosphate buffer (pH 7.25 with \( \text{KH}_2\text{PO}_4 : \text{K}_2\text{HPO}_4 = 0.3 : 0.7 \)), and different salts and salt concentrations were prepared from the stock polymer and salt solutions, potassium phosphate buffer, and deionized water. The phase systems were allowed to equilibrate and separate at room temperature (23°C) for three hours then centrifuged at 10000 g for 2 minutes to ensure complete separation of phases.

The dielectric constant, \( \varepsilon \), of phase systems with different salts was determined by measuring the refractive index, using Snell's law (Matveyev, 1966):

\[
\frac{n_1}{n_2} = \frac{\sqrt{e_1}}{\sqrt{e_2}}
\]

(3)

where \( n \) is the refractive index of a solution.

**Protein partition coefficient measurements**

For the same protein, two different concentration levels of stock protein solutions (ca. around 0.8 mg/mL and 0.5 mg/mL) were prepared and each was dialyzed against 0.01 M potassium phosphate buffer (pH 7.25). Bulk PEG, dextran, and salt solutions were weighed into polystyrene tubes and 1 g stock protein solution was added. Water was added to a final weight of 8 g, giving an overall composition of 8.5\% PEG, 8\% dextran, in 0.01 M potassium phosphate buffer at a series of added salt concentrations. Blank phase systems omitted the
protein. The systems were mixed by vortexing for 10 seconds, allowed to equilibrate for three hours, and then were centrifuged at 10000 g for 2 minutes. Top phase samples were collected using a Pasteur pipette while those from the bottom phase were obtained by piercing the bottom of the tube and allowing the lower phase to drain. The protein concentration of each phase was determined. All partitioning experiments were performed in triplicate and at room temperature, 23°C.

Potential measurements

Electrodes made of a 1.5 mm glass capillary filled with 2% agar in 3 M KCl were used to measure potential difference as described by Luther (1994). At least three measurements of each phase system were made. All potential differences reported are expressed as bottom phase minus top phase.

Flame emission adsorption measurements

Sodium and potassium ion concentrations of individual phases were determined by flame emission spectrophotometry using IL343 Flame Photometer (Instrumentation Laboratory, Inc., Lexington, Mass.). The phase solutions were diluted with a 15 mM lithium chloride solution to bring the solution concentration level to the flame emission adsorption measuring range. All the measurements were performed in triplicate and at room temperature, 23°C.
RESULTS AND DISCUSSION

Phase compositions and tie-line lengths of PEG-dextran phase systems with addition of NaCl, KCl, K₂SO₄ and potassium phosphate (pH 7.25) were obtained for the lowest and highest concentrations of added salts used in the partitioning experiments. At the 0.05 M level of added salt, the effect of the 0.01 M phosphate buffer used in each aqueous system may not be negligible. All other cases meet or exceed the criteria that addition of salt with concentration larger than ten times that of the buffer salt will suppress the influence of buffer salt (Albertsson, 1986). The addition of salts alters phase compositions and tie-line lengths (Table 2) with the concentrations of K₂SO₄ and potassium phosphate having a larger influence on tie-line lengths than do NaCl and KCl. Zaslavsky (1987, 1995) also reported the concentration effect of different salts on PEG-dextran systems and indicated that the increase of K₂SO₄ concentration reduces the total polymer concentration required for phase separation rapidly while the increase of KCl concentration has little influence.

The different salt and salt concentrations also change the measured potential difference between the phases (Table 4). These potential differences, measured by two Ag/AgCl electrodes, are the quasi-electrostatic potential differences and serve as an estimate of the electrostatic potential differences between the two phases (Albertsson, 1986; Haynes et al., 1991). Larger magnitude potential differences were observed for phase systems with potassium phosphate and K₂SO₄. While a higher salt concentration results in a smaller potential difference for all four salts, increasing concentration seems to have a larger
influence on phase systems with NaCl and KCl. Again, the 0.05 M results may be affected by the buffer salt.

Table 3 lists the partition coefficients of the four salts for these same conditions. The anions determine the partitioning of the three potassium salts as has been reported by others (Zaslavsky, 1995; Johansson et al., 1973; Hartounian et al., 1994) with the order of partition coefficients as chloride salts > sulfate salts > phosphate salts, where the phosphate salts have the most uneven distribution between the two phases. Additionally, for chloride salts, the extent of partitioning decreases with the increase of salt concentration, while for sulfate and phosphate salts, the extent of partitioning increases. The same effect of phosphate salt concentration was reported by Bamberger et al. (1984). These effects on salt distribution are the result of changes in solvent structure, and polymer-salt and polymer-polymer interactions. Zaslavsky (1995) reported that increasing the concentration of water-structure-making ions such as sulfate and phosphate ions will reduce PEG solubility in the dextran phase and salt solubility in the PEG phase, which leads to a more uneven distribution of polymers and salts between the two phases. This is also the case for PEG-salt systems as increase of the salt concentration will push more salts from the PEG phase to the salt phase (Hartounian et al., 1994). On the other hand, the increase of water-structure-breaking ions concentration like chloride ions concentration will increase the polymer compatibility (Brooks et al., 1984; Zaslavsky, 1995), which means more PEG transferring from top phase to bottom and a more even salt distribution between the two phases. The results in Table 2, the salt effects on phase compositions and tie-line lengths, and the results in Table 3, the salt distributions are
related each other. Zaslavsky et al. (1988) reported that for phase systems with the same salt and salt concentration but different polymer concentrations:

\[ \ln K_s \propto \Delta C_{\text{polymer}} \]  

(4)

where \( K_s \) is the salt partition coefficient and \( \Delta C_{\text{polymer}} \) is PEG concentration difference between the two phases, which is in turn, often proportional to the tie-line length. The same trend predicted from Equation 4 is evident in our salt partitioning and tie-line length results (Table 2 and 3); however, in our case, the overall polymer concentrations has remained the same and phase compositions were changed indirectly by the addition of different salts with different concentrations.

As a result of the salt changes, a protein will encounter differences in both polymer and salt compositions of the phases as well as different interfacial potential differences among the phase systems. Figures 1a - d show the resulting partition coefficients of the six T4 lysozymes in all these phase systems. These results not only illustrate the influence of different salts and salt concentrations on the protein partitioning, but also indicate the effects of surface charge and charge distribution of the protein on partitioning. The importance of protein surface charge is evident in the different partitioning behaviors among the six T4 lysozymes. Yet surface charge alone is not a sufficient explanation as seen in the differing partition coefficients of U1 and K16/136E and U2 and K16E in all phases systems even though these pairs have matching surface charge. As discussed previously (Fan et al., 1997), the corresponding point mutation and fusion-tail versions differ in charge polarity leading to differences in dipole-dipole and charge-dipole interactions.
One general trend from Figures 1a - d is that the addition of any of the salts increases the partition coefficient of all the proteins (though the "increase" in these cases means a more equal distribution). The increase of the partition coefficient of positively charged proteins with the increase of salt concentrations was also reported by Albertsson (1986). However, the magnitude of the change as well as the absolute value of the partition coefficient depends on the salt. First, we can see that the protein net charge has a more significant impact on partitioning in phase systems with $K_2SO_4$ and potassium phosphate as the differences in partition coefficients among the charge mutants are greater than those observed in phase systems with NaCl and KCl.

The larger influence of protein charge for systems with $K_2SO_4$ and potassium phosphate is more easily interpreted when replotted according to Equation 2 in Figures 2 a-d. The slopes of these plots indicates that the charge effects reflect the magnitudes of the salt-determined potential differences. Table 4 includes the potential differences calculated from the slopes using Equation 2. While the quantitative agreement with the measured values is not good, qualitatively the larger potentials of phase systems with $K_2SO_4$ and potassium phosphate are confirmed.

Table 4 also shows that, as a general trend, the potential differences decrease with increased salt concentration. Phase systems with $K_2SO_4$ and potassium phosphate present a relatively stable electrostatic effect as the potential differences change less significantly with the increase of salt concentrations. For phase systems with NaCl and KCl, we can even observe a change of direction of the potential difference with the increase of salt.
concentrations; however, this is probably the result of the influence of the phosphate buffer at low added salt concentration.

From the intercepts of the linear relationships in Figure 2a-d, we can observe the salt effects on non-electrostatic interactions, which are included in the $K_Q$ term of Equation 2. Figure 2a and 2b show that for phase systems with NaCl and KCl, different salt concentrations result in small changes of intercept, which means the non-electrostatic interactions under different salt concentrations are relatively constant, and the addition of NaCl and KCl affects the protein partitioning mainly through altering the potential difference between the two phases. In contrast, for systems with K$_2$SO$_4$ and potassium phosphate. Figure 2c and 2d show that changing salt concentration has smaller effect on the slope but does affect the intercept significantly, which implies that the significant influence of salt concentration on the protein partition coefficients results from non-electrostatic effects. This is consistent with our earlier observation that the increase of K$_2$SO$_4$ and potassium phosphate concentration has more significant effects on phase compositions, tie-line lengths and salt partitioning than that of NaCl and KCl as shown in Table 2 and Table 3. However, this non-electrostatic effect cannot be explained by the tie-line effect accounted for by Equation 1. It predicts that a larger tie-line length will result in a lower (more negative) intercept; however, here higher salt concentration gives a larger tie-line length but a larger (less negative) intercept. This implies that the change of salt concentration has also altered the parameter $a$ in Equation 1, indicating altered protein-polymer molecular interactions.
Therefore, the salt effects on charged protein partitioning can be viewed as a result of the interplay of the electrostatic effects through the salt effects on phase potential differences, and non-electrostatic effects through the salt effects on phase polymers, salt compositions and protein-polymer molecular interactions. Equation 2 seems proper in qualitatively explaining the partitioning results of proteins with different surface charge in these systems. Zaslavsky (1995) countered with the alternative explanation that salt effects on protein-polymer-salt molecular interactions are caused by different dielectric properties between the two phases resulting from salt effects on solvent structure, and salt and polymer compositions. However, Table 5, while indicating that there is a small dielectric constant difference between the top and bottom phases of PEG-dextran systems with different concentrations of NaCl, lists the results, shows that the influence of salt concentration is less than 0.5%, i.e. almost negligible. The same results were found for phase systems with KCl, K$_2$SO$_4$ and potassium phosphate. The results may imply that the measured dielectric constant is the dielectric property of bulk phases and has less influence on molecular interactions. As salt effects on different molecular interactions between polymers, proteins and salts take place around protein and polymer molecular surfaces and the hydration layers, dielectric properties at the surface or around the vicinity of these molecules are more important.
CONCLUSIONS

We have obtained the partitioning results of salts and T4 lysozymes in PEG-dextran phase systems with different salts and salt concentrations. Potassium and sulfate salts and chloride salts display different partitioning trends with the increase of salt concentration. While for all the T4 lysozymes, increase in the salt concentration will push more protein from bottom phase to top phase, T4 lysozyme fusion tail mutants display different partitioning behavior from those of point mutation mutants with matching surface charge. The partitioning results of T4 lysozyme point mutation mutants provided us the protein charge effects in phase systems with different salts and salt concentrations under the same pH. By using Equation 2, we were able to find how different salts and salt concentrations would bring changes of electrostatic effect, i.e. the potential differences, and non-electrostatic effects, i.e. the $K_0$ term, and how these changes affect the charged protein partitioning. In phase systems with addition of potassium phosphate and $K_2SO_4$, a larger magnitude of the potential difference between the two phases was found. In phase systems with NaCl and KCl, different salt concentrations affect the protein partitioning largely through its influence on the potential difference. These partitioning results provide us hints on how to choose the right salt and salt concentration in order to select the target protein using protein surface charge in aqueous two-phase extraction.
ACKNOWLEDGMENT

We thank Dr. Brian Matthews for providing the cultures to produce the T4 lysozyme and mutants with charge-change point mutations and Dr. Clark Ford of Iowa State University for guidance and facilities in preparing the genetic constructions. This material is based upon work supported by the National Science Foundation under Grant No. BCS-9108583. The government has certain rights in this material. Weiyu Fan was partially supported by the Biotechnology Byproducts Consortium under a contract from the U.S. Department of Agriculture.

REFERENCES


Table I. Net charge of t4 lysozyme mutants and fusions at pH=7.25*

<table>
<thead>
<tr>
<th>T4 Lysozyme</th>
<th>Wild</th>
<th>Single</th>
<th>Double</th>
<th>Triple</th>
<th>U1</th>
<th>U2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Mutant</td>
<td>Mutant</td>
<td>Mutant</td>
<td>Fusion</td>
<td>Fusion</td>
<td></td>
</tr>
<tr>
<td>Net charge</td>
<td>9.84</td>
<td>7.84</td>
<td>5.84</td>
<td>3.84</td>
<td>5.84</td>
<td>7.84</td>
</tr>
</tbody>
</table>

*Charge estimated by using a combination of titration, isoelectric focusing, and calculation from the Henderson-Hasselbalch relationship and pK values from Lehninger (1982) (Luther and Glatz, 1994).

Table II. The phase compositions of PEG-dextran systems with salts at the lowest and highest salt concentrations used

<table>
<thead>
<tr>
<th></th>
<th>NaCl</th>
<th>KCl</th>
<th>K_2SO_4</th>
<th>Potassium Phosphate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05 M</td>
<td>0.5 M</td>
<td>0.05 M</td>
<td>0.5 M</td>
</tr>
<tr>
<td>Top Phase</td>
<td>(12, 1.6)</td>
<td>(11, 1.4)</td>
<td>(16.33)</td>
<td>(17.29)</td>
</tr>
<tr>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
</tr>
<tr>
<td>Bottom Phase</td>
<td>(22, 33)</td>
<td>(24, 33)</td>
<td>(41.30)</td>
<td>(4.2.29)</td>
</tr>
<tr>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
<td>(PEG, Dextran) %</td>
</tr>
<tr>
<td>Tie-Line Length %</td>
<td>32.9</td>
<td>33.0</td>
<td>29.3</td>
<td>29.1</td>
</tr>
</tbody>
</table>
Table III. The partition coefficients of different salts in PEG-dextran systems with different concentration level

<table>
<thead>
<tr>
<th>Salt Concentration</th>
<th>NaCl</th>
<th>KCl</th>
<th>K₂SO₄</th>
<th>Potassium Phosphate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05 M</td>
<td>0.90</td>
<td>0.85</td>
<td>0.74</td>
<td>0.68</td>
</tr>
<tr>
<td>0.1 M</td>
<td>0.93</td>
<td>0.91</td>
<td>0.70</td>
<td>0.66</td>
</tr>
<tr>
<td>0.2 M</td>
<td>0.97</td>
<td>0.93</td>
<td>0.67</td>
<td>0.59</td>
</tr>
<tr>
<td>0.4 M</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.49</td>
</tr>
<tr>
<td>0.5 M</td>
<td>0.98</td>
<td>0.96</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table IV. Potential differences (Top-Bottom) calculated from the partition data of T4 lysozymes point mutants using the linear relationship of Equation 2 (Cal) (Figure 2a-d), and potential differences (Top-Bottom) measured from Ag/AgCl electrodes (Exp) *

<table>
<thead>
<tr>
<th>Salt</th>
<th>NaCl</th>
<th>KCl</th>
<th>K$_2$SO$_4$</th>
<th>Potassium Phosphate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cal (mV)</td>
<td>Exp (mV)</td>
<td>Cal (mV)</td>
<td>Exp (mV)</td>
</tr>
<tr>
<td>0.05 M</td>
<td>1.5 ±0.04</td>
<td>0.52 ±0.04</td>
<td>0.20 ±0.03</td>
<td>1.7 ±0.06</td>
</tr>
<tr>
<td>0.1 M</td>
<td>0.01 ±0.03</td>
<td>0.04 ±0.03</td>
<td>0.34 ±0.04</td>
<td>1.0 ±0.05</td>
</tr>
<tr>
<td>0.2 M</td>
<td>-0.46 ±0.04</td>
<td>0.25 ±0.04</td>
<td>0.85 ±0.04</td>
<td>0.95 ±0.04</td>
</tr>
<tr>
<td>0.4 M</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.0 ±0.03</td>
</tr>
<tr>
<td>0.5 M</td>
<td>-0.90 ±0.03</td>
<td>0.12 ±0.03</td>
<td>0.74 ±0.02</td>
<td>-</td>
</tr>
</tbody>
</table>

* ± standard deviation of mean

Table V. The Dielectric constant of top and bottom phase PEG-dextran systems with different NaCl concentrations*

<table>
<thead>
<tr>
<th>Concentration</th>
<th>Top</th>
<th>Bottom</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05 M</td>
<td>81.47±0.02</td>
<td>83.37±0.01</td>
</tr>
<tr>
<td>0.1 M</td>
<td>81.51±0.01</td>
<td>83.44±0.02</td>
</tr>
<tr>
<td>0.2 M</td>
<td>81.63±0.02</td>
<td>83.55±0.03</td>
</tr>
<tr>
<td>0.5 M</td>
<td>82.02±0.02</td>
<td>83.85±0.01</td>
</tr>
</tbody>
</table>

*± standard deviation of mean
FIGURE LEGEND

Figure 1. Partition Coefficient of all six T4 lysozymes in the PEG3350/dextran-40 aqueous two-phase systems, at pH 7.25 (0.01M potassium phosphate buffer), with addition of different salts: (a) NaCl, (b) KCl, (c) K₂SO₄ and (d) potassium phosphate at pH 7.25 with KH₂PO₄ : K₂HPO₄ = 0.3 : 0.7. The partition results at 0.05 M salt concentration could be influenced from both added salts and 0.01M phosphate buffer salts.

Figure 2. Partition Coefficient of T4 lysozyme Wild Type, K/16E, K/16/135E, K16/135/147E versus protein net surface charge in phase systems with addition of (a) NaCl, (b) KCl, (c) K₂SO₄ and (d) potassium phosphate at pH 7.25 with KH₂PO₄ : K₂HPO₄ = 0.3 : 0.7.
(Figure 1a)
(Figure 1b)
(Figure 1c)
(Figure 1d)
Protein Net Charge

(Figure 2a)
(Figure 2b)
Protein Net Charge

(Figure 2c)
(Figure 2d)
CHAPTER 5. CHARGE-CHANGE MUTATION CONTRIBUTION TO THE OSMOTIC SECOND VIRIAL COEFFICIENT FOR GENETIC ENGINEERED PROTEINS

A paper to be submitted to Biotechnology & Bioengineering

Weiyu Fan and Charles E. Glatz
Department of Chemical Engineering, Iowa State University, Ames, IA 50011

ABSTRACT

The osmotic second virial coefficient is often used to characterize the thermodynamic properties of dilute protein solutions. Comparison of the second virial coefficient of wild type to those of mutants can lead to an understanding of how changes of protein surface properties can affect the thermodynamic properties of proteins in solutions.

We obtained three forms of T4 lysozyme: (1) a double point mutation (Lys→Glu, DBL), (2) an Arg₂ fusion (U1) to a triple Lys→Glu point mutation, and (3) the wild type (WT). The two mutants have matching net charge but different polarity. The second virial coefficients of DBL, U1 and WT in solutions of different ionic strengths were measured by multi-angle laser light scattering. The second virial coefficients were also calculated from McMillan-Mayer theory using estimates of the molecular interaction forces of the three forms. The results from laser light scattering agreed with the theoretical calculations at higher but not lower ionic strengths. Comparison of the second virial coefficients of WT to
those of the point and fusion-tail mutants reveals the effects of surface charge and dipole moment of the protein on electrostatic intermolecular interactions.

INTRODUCTION

Genetic engineering enables manipulation of protein surface properties to enhance the yield and selectively of downstream processes, such as precipitation, aqueous two-phase extraction, and chromatography. Some successful applications of genetic engineering technology in downstream processing have been reported (Ford et al., 1991; Kohler et al., 1991; Niederauer and Glatz, 1992; Luther and Glatz, 1994). Unfortunately, lack of understanding of how protein surface changes alter solution thermodynamics limits prediction of protein solubility, partitioning, and binding in such processes.

The osmotic second virial coefficient is often used to describe the thermodynamic properties of a protein solution. It is one of a series of parameters in the virial expansion model for osmotic pressure (Hill, 1962):

\[
\frac{\pi}{kT} = c + B_2 c^2 + B_3 c^3 + \cdots
\]

where \(\pi\) is the osmotic pressure; \(B_n\) is the nth virial coefficient and has dimension \((m^3/molecule)^{n-1}\); \(c\) is the solute number density in unit of molecules/m^3. According to McMillan-Mayer theory, \(B_n\) represents the interaction of \(n\) bodies in the solution. The second virial coefficient \(B_2\), a measure of the two-body interaction, can be expressed as an integral of the potential of mean force, \(w(r, T)\), between the two molecules:
$B_2 = -\frac{1}{2} \int_{d_2 + 3 \sigma}^{\infty} \left\{ \exp\left[ -\frac{w(r, T)}{kT} \right] - 1 \right\} 4\pi r^2 dr \tag{2}$

where $d_2$ is the protein diameter, $r$ is the distance between the two molecules, and the lower limit is taken as $d_2 + 3 \sigma$ to account for a layer of water bound to the protein (Vilker et al., 1981). The second virial coefficient estimate of the molecular interaction force has been used to predict properties of protein solutions, such as stability and osmotic pressure (Vilker et al., 1981; Coen et al., 1995; Pjura et al., 1995; Antipova and Semenova, 1995). A comparison of the second virial coefficient of wild-type protein to those of mutants can lead to an understanding of how change of protein surface properties by genetic manipulations can affect the protein molecular interactions in a solution as well as the properties of the protein solution.

Multi-angle laser light scattering can be used to measure the protein second virial coefficient (King et al., 1988; Rathbone et al., 1990). The measurement is based on a graphical technique, the Zimm plot, to extrapolate and obtain the values of the zero angle Rayleigh ratio at different protein concentration. The second virial coefficients, $B_{ii}$, and molecular weight, $M_i$, can be obtained directly from the slope and intercept of the plot at zero angle (Kratochvil, 1987):

$$\frac{K' c_i}{R_0} = \frac{1}{M_i} + 2B_{ii} c_i \tag{3}$$
where $R_g$ is the Rayleigh ratio, directly related to the scattered light intensities at angle $\theta$; $c$, is the concentration of the solute in g/mL; and $K^*$ is the single-component optical constant.

A model calculating the second virial coefficient based on Equation 2 is also desirable in order to identify how different molecular forces contribute to the potential of mean interaction force. According to Vilker et al. (1981), the potential of mean force, $w(r, T)$, for a protein diluted in an aqueous salt solution can be evaluated as a sum of potentials of the following interactions:

$$w(r, T) = w_{hs} + w_{q-q} + w_{q-\mu} + w_{\mu-\mu} + w_d$$

where $r$ is the distance from the center of the protein molecule and $T$ is the temperature; the terms on the right hand side represent, respectively hard-sphere, charge-charge, charge-dipole, dipole-dipole, and dispersion / van de Waals interactions. Table 1 lists the expressions for the potentials of the various interactions.

Comparison of virial coefficients of proteins differing only in net charge or dipole moment would provide insight on how different electrostatic interactions, such as charge-charge and charge-dipole interactions, caused by different surface charge and charge distribution, affect the potential of mean force (Wills et al., 1995; Kozak et al., 1968; Cole. 1996). Ideally, identification of the dominate interactions can guide genetic engineering strategies to manipulate separation behavior.
MATERIALS AND METHODS

Proteins

WT, DBL and U1 T4 lysozyme were produced from the expression vector pHN1403, which has been transformed into *E. coli* strain RR1. DBL, the double mutant, has been modified through site-directed mutagenesis to replace two lysine residues (K16 and K135) with glutamic acids (Dao-Pin, 1991, strains provided by B.W. Matthews, University of Oregon). U1 is constructed with a fusion of two arginines tail to the carboxyl terminus of a T4 lysozyme triple mutant, with three lysine residues (K16, K135 and K147) being replaced by glutamic acids (Fan et al., 1997). DBL and U1 have the same net charge, which differs from that of WT, but different polarity. Full enzymatic activity is observed for all three T4 lysozymes.

Protein production and activity assay

T4 lysozyme production was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) and the enzymes were purified as described by Luther and Glatz (1994). Lysozyme activity was measured by following the clearing of a *Micrococcus lysodeiktkus* cell suspension versus time (Luther and Glatz, 1994).

Multi-angle laser light scattering

Light scattering measurement were made using the DAWN-B light scattering instrument (Wyatt Technologies, Santa Barbara, CA), a batch system employing a laser light beam at 632.8 nm, and measuring the intensity of light scattered at 15 different angles ranging from 23° to 128°. Refractive index increments were measured on an OPTILAB 903
Differential Refractometer (Wyatt Technologies, Santa Barbara, CA) also operating at 632.8 nm. All the buffers were prepared using deionized water (Barnstead, NanopureII water treatment system, Barnstead CO., Boston, MA) and then filtered (0.2 μm, Gelman Sciences, Ann Arbor, MI). Protein samples were dialyzed overnight against potassium phosphate buffer (0.01M, pH 7.25) and filtered (0.02 μm Anotop syringe tip filters, Whatman). All experiments were performed at room temperature (23°C).

The glass scintillation vials used as sample cells in the DAWN-B instrument required about 6 ml sample to fill the light path. The intensity of scattered light was collected first for the most concentrated protein sample. The protein solution in the vial was then diluted by adding filtered (0.02 μm Anotop syringe tip filters, Whatman) dialysis buffer. For each protein, more than 8 concentration levels were usually measured by such serial dilutions.

The DAWN 1.0 version software (Wyatt Technologies, Santa Barbara, CA) was used to handle the collected data. The software converts the data to a Zimm plot automatically and calculates the value of the second virial coefficient and molecular weight from Equation 3 after allowing the user to review the goodness of fit to the data.

RESULTS AND DISCUSSION

The crystallographic data of wild-type T4 lysozyme were obtained from the Protein Data Bank (PDB) (Brookhaven National Laboratory, Brookhaven, New York, USA). The structural data for DBL and U1 were extrapolated from the WT structure by using the Swiss-Model (an automated protein modeling server running at the Geneva Biomedical Research
Institute. Glaxo Wellcome Research and Development, S.A., Switzerland). Table 2 reports the molecular properties of the three T4 lysozymes. Software SURVOL (from Pontius. J., UCMB, l'Universite Libre de Bruxelles, Brussels, Belgium) was used to analyze the volume and surface area of the enzymes. The equivalent spherical radius was estimated from the accessible molecular volume. The estimation of charge of T4 lysozymes has been reported previously (Luther and Glatz, 1994). U1 has the same charge as that of DBL. Dipole moments were calculated from PDB files of crystallographic data using the program GRASP (Graphical Representation and Analysis of Surface Properties, a program running under UNIX system, Anthony Nicholls, 1992). The results indicate that because of the different charge distribution, U1 and DBL have different dipole moments.

By using the protein properties provided in Table 2 and the potential expressions in Table 1, the potentials of the various interaction forces can be calculated. It can be foreseen that WT will have larger charge-charge, charge-dipole, and dipole-dipole forces because of its larger charge number and dipole moment. Compared to U1, DBL will have larger dipole-dipole and charge-dipole forces because of DBL’s larger dipole moment. Figures 1a, b, and c show the calculated results for the magnitude of the potentials of different forces around the vicinity of the protein molecules in the solution with \( I=0.01 \) M potassium phosphate. The potential energy is normalized by \( kT \) and center-center distance is normalized by the equivalent spherical radius \( a \). The repulsive charge-charge force has the largest influence for each protein and therefore, WT has the largest repulsive interaction force. Charge-dipole and dispersion are the two most important attractive forces. DBL has large dipole-dipole and
charge-dipole attractions because of the larger dipole moment. The charge-charge and
dispersion forces are the longer range forces as their magnitudes remain fairly large as the
separation distance \( r \), increases.

Figures 2a, b, and c show the effects of ionic strength on the protein molecular
interaction forces. Ionic strengths alters the Debye lengths (Hill, 1962) as

\[
\kappa^{-1} = \frac{3.04}{\sqrt{I}}
\]

(5)

where \( I \) is ionic strength and \( \kappa^{-1} \) is the Debye length in \( \AA \). A smaller Debye length means a
stronger double layer screening of the electrostatic interaction forces, and shorter working
distance for these forces. Hence, for all three proteins, because the charge-charge interaction
is the most important one, larger ionic strength results in a less positive, or less repulsive
overall interaction force. For the same reason, the overall potentials of interaction at \( I = 0.1 \)
\( \text{M} \) differ less among the three proteins.

The second virial coefficients of all the three lysozymes at different ionic strengths
were determined from laser light scattering (Table 3). Qualitatively, the results are in good
agreement with the theoretical calculations. WT has the largest second virial coefficient in
all the solutions because of the largest charge-charge repulsive force. The screening effects
of ionic strength are also reflected in the lower values of \( B_2 \) at higher ionic strengths.
Comparing to U1, because of larger charge-dipole and dipole-dipole attractions, DBL has a
lower value of \( B_2 \).
Figures 3a, b, and c compare the results from experiments to those calculated from Equation 2. The Simpson method was used in computing the integration. The potentials of mean force were calculated according to Table 1. An accurate Hamaker constant for T4 lysozyme is not available, but since all globular proteins should have similar Hamaker constants (Cantor and Schimmel, 1980), a Hamaker constant of 1 kT was used for all three T4 lysozymes in evaluating the potential of dispersion force (Nir, 1976; Coen et al., 1995). Again, it is evident that qualitatively, the model is in good agreement with the experiment. Changes of the second virial coefficient found from experiments, caused by protein surface charge, dipole moment, as well as ionic strength, have been reflected in the model. However, the figure shows that for all the three lysozymes, the values from experiments at lower ionic strength, especially, at 0.01M potassium phosphate, are an order of magnitude larger than those calculated from the model. On the other hand, at higher ionic strengths, while the model has a better prediction on the second virial coefficient, it is interesting to note that the values from the experiments are even smaller than those from the calculation.

For WT, with the largest surface charge, the second virial coefficients measured from the experiment are larger than the calculated values for I=0.01, 0.02 and 0.05 M potassium phosphate. For DBL and U1, the measured second virial coefficients are larger than the calculated values only at I=0.01 M potassium phosphate. This would indicate that the deviations derive mainly from the charge-charge interaction term. This may imply that the two-term virial expansion model used in determining the second virial coefficient from light scattering is not sufficient to describe the charge-charge interactions at low ionic strengths.
Charge-charge forces are long-range interaction forces and a low ionic strength means a larger Debye length, i.e. a longer charge-charge force working distance (Table 4). Hence, at low ionic strengths, charge-charge interactions from a third or even a fourth molecule might also contribute significantly to the potential of mean force, which brings the need to include the third or even fourth virial coefficient into the expansion. Cole (1996) reported the same discrepancy in the second virial coefficient of hCMV after comparing experimental values from velocity sedimentation measurements to those calculated on the basis of excluded volume and electrostatic effects.

At higher ionic strengths, or with lower surface charge, we found that the calculated values of the second virial coefficients become lower than those from experiments. As we have analyzed above, the dispersion force is the long range force and is the most important attractive force. Furthermore, since the Hamaker constant we used is a tentative one, it is very possible that the underestimation of the dispersion force is the main reason for the difference at higher ionic strengths.

One can use the measured values of the second virial coefficients to estimate the Hamaker constant, by trying different values of Hamaker constant and fitting the calculation results to those from experiment. Using this method, the Hamaker constants, $A/kT$, of WT, DBL and U1 at $I=0.01$ M potassium phosphate are 10.4, 12.7 and 8.1 respectively, rather than the value of 1 assumed in Figures 1 and 3. The similar values are expected as the three lysozymes are the same proteins except for surface charge and dipole moment. However, the
large values of Hamaker constant may indicate that there are other attractive forces that we did not include in our calculations (Coen et al., 1995).

CONCLUSIONS

We have measured, by laser light scattering, the second virial coefficients of wild-type T4 lysozyme, the DBL double mutant and the U1 fusion. A model of the McMillan-Mayer type, based on excluded volume, charge-charge, charge-dipole, dipole-dipole, charge-fluctuation, and dispersion interactions between two protein molecules in solution was used to calculate the second virial coefficient. Both the experimental measurements and modeling calculations imply that the charge-charge interactions caused by the protein surface charge are the major contribution to the potential of mean force. Therefore, higher surface charge (WT) results in a larger second virial coefficient. Dipole attractions can not be neglected as DBL and U1 have different second virial coefficients due to the different dipole moments. The influence of ionic strength is mainly represented in the screening effects of the ion atmosphere. The discrepancy of the second virial coefficient between results from experiments and from calculations indicate that at low ionic strengths, a higher order of virial expansion is needed, while at high ionic strengths, it is necessary to obtain a better evaluation of the Hamaker constant of the proteins.
ACKNOWLEDGMENT

We thank Dr. Brian Matthews for providing the cultures to produce the T4 lysozyme and mutants with charge-change point mutations and Dr. Clark Ford and Dr. Ufuk Bakir of Iowa State University for their part in preparing the genetic construct. This material is based upon work supported by the National Science Foundation under Grant No. BCS-9108583. The government has certain rights in this material. Weiyu Fan was partially supported by the Biotechnology Byproducts Consortium under a contract from the U.S. Department of Agriculture.

REFERENCES


### Table I. Intermolecular potential function.

<table>
<thead>
<tr>
<th>Type</th>
<th>Potential Function</th>
<th>Screening factor</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excluded volume, $w_{hv}$</td>
<td>$4v_m$</td>
<td></td>
<td>Vilker et al., 1981</td>
</tr>
<tr>
<td>Charge-charge, $w_{q-q}$</td>
<td>$\frac{(Ze)^2}{4\pi e r}$</td>
<td>$e^{-\kappa(r-\alpha)}$</td>
<td>Vilker et al., 1981</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(1+\kappa\alpha)^2$</td>
<td></td>
</tr>
<tr>
<td>Charge-dipole, $w_{q-\mu}$</td>
<td>$\frac{2(ze)^2 \mu^2}{3 e^2 kT \alpha^4}$</td>
<td></td>
<td>Phillips, 1974</td>
</tr>
<tr>
<td></td>
<td>$\left{ \frac{3(1+\kappa\alpha)e^{-\kappa(\alpha-\alpha)}}{(1+\kappa\alpha)(2+2\kappa\alpha+(\kappa\alpha)^2 + (1+\kappa\alpha)e/\varepsilon)^2} \right}^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dipole-dipole, $w_{\mu-\mu}$</td>
<td>$\frac{2 \mu^4}{3 kT \alpha^2 \varepsilon^6}$</td>
<td></td>
<td>Phillips, 1974</td>
</tr>
<tr>
<td></td>
<td>$\left{ \frac{3[2+2\kappa\varepsilon+(\kappa\varepsilon)^2]e^{-\kappa(r-\alpha)}}{2+2\kappa\varepsilon+(\kappa\varepsilon)^2 + (1+\kappa\varepsilon)e/\varepsilon} \right}^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion, $w_d$</td>
<td>$\frac{A}{6}\left(\frac{2}{s^2} + \frac{2}{s^1 - 4} + \ln(\frac{s^1 - 4}{s^1})\right)$</td>
<td></td>
<td>Coen et al., 1995, Nir, 1976, Vilker et al., 1981</td>
</tr>
<tr>
<td></td>
<td></td>
<td>where $s = r/\alpha$ and $A$ is the Hamaker constant</td>
<td></td>
</tr>
</tbody>
</table>

where $\kappa$ is the inverse of the Debye length and $\alpha$ is the equivalent spherical radius

where $\varepsilon_r$ is the dielectric constant at the protein molecular surface and $\varepsilon$ is the solution dielectric constant
Table II. Molecular properties of the three T4 lysozymes.

<table>
<thead>
<tr>
<th>T4 Lysozyme</th>
<th>Wild Type</th>
<th>DBL</th>
<th>UI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular weight</td>
<td>18400</td>
<td>18400</td>
<td>18800</td>
</tr>
<tr>
<td>Molecular volume, $A^3$</td>
<td>24713</td>
<td>24713</td>
<td>24713</td>
</tr>
<tr>
<td>Equivalent spherical radius, $a$, Å</td>
<td>20.9</td>
<td>20.9</td>
<td>20.9</td>
</tr>
<tr>
<td>Surface Charge</td>
<td>9.8</td>
<td>5.8</td>
<td>5.8</td>
</tr>
<tr>
<td>Dipole Moment, $\mu$</td>
<td>359</td>
<td>290</td>
<td>245</td>
</tr>
</tbody>
</table>

Table III. The second virial coefficients from laser light scattering*.

<table>
<thead>
<tr>
<th>The Second Virial Coefficient</th>
<th>Wild Type</th>
<th>DBL</th>
<th>UI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_2 \times 10^4$ (mL·mol/g²)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01 M Potassium Phosphate</td>
<td>120±20</td>
<td>40±10</td>
<td>98±20</td>
</tr>
<tr>
<td>0.02 M Potassium Phosphate</td>
<td>44±10</td>
<td>3.5±0.6</td>
<td>3.9±1.5</td>
</tr>
<tr>
<td>0.05 M Potassium Phosphate</td>
<td>25±6</td>
<td>-0.64±0.1</td>
<td>2.4±0.4</td>
</tr>
<tr>
<td>0.1 M Potassium Phosphate</td>
<td>2.0±0.5</td>
<td>-3.9±0.6</td>
<td>-1.3±0.2</td>
</tr>
</tbody>
</table>

* Standard deviation of the three replicates; each value obtained from an individual measurement by curve fitting of the Zimm plot has standard deviation less than 25%.
**Table IV.** Comparisons of protein molecular distance and Debye length

<table>
<thead>
<tr>
<th>Salt Concentration</th>
<th>0.01M</th>
<th>0.02M</th>
<th>0.05M</th>
<th>0.1M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average protein-protein inter-molecular distance ($Å$)*</td>
<td>249</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debye length ($Å$)</td>
<td>19.6</td>
<td>13.9</td>
<td>8.78</td>
<td>6.21</td>
</tr>
<tr>
<td>Charge-charge force working distance ($Å$)*</td>
<td>152</td>
<td>122</td>
<td>95</td>
<td>81</td>
</tr>
</tbody>
</table>

* Calculation based on protein concentration at 2 mg/mL.

* Calculation based on the charge-charge force of Double Mutant; the working distance is determined when at $r = \text{working distance}$, the potential of the charge-charge force $< 10^{-3}$ of the potential of the charge-charge force at $r = 2a$, where $a$ is the radius a protein molecule.
FIGURE LEGENDS

Figure 1a, 1b, 1c. Magnitude of screened-pair potential energies of interactions at I=0.01 M potassium phosphate. (a) Wild Type; (b) DBL, double mutant; (c) U1, fusion mutant.

Figure 2a, 2b, 2c. Pair potential of the sum of different interaction forces at different ionic strengths. (a) Wild Type; (b) DBL, double mutant; (c) U1, fusion mutant.

Figure 3a, 3b, 3c. Comparisons of the second virial coefficients at different ionic strengths from laser light scattering to those calculated from the model. (a) Wild Type; (b) DBL, double mutant; (c) U1, fusion mutant.
\[ \text{charge-charge} \]
\[ \text{charge-dipole} \]
\[ \text{dipole-dipole} \]
\[ \text{dispersion} \]

(Figure 1a)
(Figure 1c)
(Figure 2a)
Figure 2b

Graph showing the relationship between \( \frac{w}{MT} \) and \( r/a \) for different concentrations: 
- \( l=0.01 \text{ M} \)
- \( l=0.02 \text{ M} \)
- \( l=0.05 \text{ M} \)
- \( l=0.1 \text{ M} \)
(Figure 3a)
• From Experiment
• From Calculation

(Figure 3b)
From Experiment

From Calculation

(Figure 3c)
CHAPTER 6. GENERAL CONCLUSIONS

We have studied the partitioning behavior of wild type T4 lysozyme and its charge-change mutants in PEG-dextran phase systems with different overall polymer concentrations, or the same overall polymer concentration but with addition of different salts and salt concentrations. In almost all the phase systems, T4 lysozyme with different charge displays different partitioning, which indicates the importance of this charge effect and feasibility of improving the selectivity of the target protein in aqueous two-phase extraction by changing the protein surface charge.

The potential differences between the two phases of all the phase systems were measured by two Ag/AgCl electrodes. These potential differences, $\Delta \Phi$, are the quasi-electrostatic potential differences and serve as an estimate of the electrostatic potential difference between the two phases, $\Delta \phi$, in Equation 2.12. We found that Equation 2.12 can qualitatively predict the protein charge effect as a larger charge effect on the partitioning of T4 lysozyme mutants was observed in phase systems with larger potential differences. A larger overall polymer concentration results in a larger potential difference between the two phases, and phase systems with the addition of phosphate or sulfate salts have larger potential differences than those with chloride salts. The influence of salt concentration on the charged protein partitioning behavior varies with the salt.

In addition to the different partitioning behavior of T4 lysozyme with different surface charges, T4 lysozyme mutants with the same charge but different charge distribution
brought by different genetic engineering methods display very different partitioning behavior in all the phase systems, which implies the importance of charge distribution as well as different molecular interactions brought by the different charge distributions. A model based on a virial expansion thermodynamic model and experimental regression was derived in trying to illustrate how the changes of protein charge and charge distribution will affect the protein-polymer molecular interactions in phase systems, and in turn, lead to different partitioning behavior. The second virial coefficient of T4 lysozymes with PEG or dextran measured from laser light scattering was used to quantify different molecular interactions caused by changes of protein charge and charge distribution. However, while the results from laser light scattering qualitatively characterized the change of molecular interactions of different T4 lysozymes, applying the measured second virial coefficient from laser light scattering to the model did not give a good prediction of the protein partition coefficient.

Finally, the second virial coefficients of T4 lysozymes measured from laser light scattering were compared to the values calculated from a model of the McMillan-Mayer type, based on excluded volume, charge-charge, charge-dipole, dipole-dipole, charge-fluctuation, and dispersion interactions between two protein molecules in solution. The results provided evidence that the charge-charge interactions caused by the protein surface charge are the major contribution to the potential of mean force and T4 lysozyme fusions have different interactions in the phase solution from those of mutants due to the different charge-dipole and dipole-dipole interactions. The results of this research indicate the change of protein surface properties will cause different molecular interactions, which in turn, cause different
partitioning behavior in aqueous two-phase systems. These results are meaningful on how to use genetic engineering methods to enhance the target protein selectivity in the downstream processing.
LITERATURE CITED


1984. Electrostatic and Electrokinetic Potentials in Two Polymer Aqueous Phase

Cabezas, H., Evans, J.D., and Szlag, D.C., 1990. in Down-stream Processing and
Bioseparations: Recovery and Purification of Biological Products. ACS Symposium

Cheng, L., Joelsson, M., Johansson, G., 1990. Combination of Polymer-Bound Charged
Groups and Affinity Ligands for Extraction of Enzymes by Partitioning in Aqueous Two-

Equilibria and Intermolecular Potentials. AIChE J., 41:996.


Mathematical Model Describing Behavior of Biomolecules in Aqueous Two-Phase

Diamond, A.D., and Hsu, J.T., 1989. Fundamental Studies of Biomolecules Partitioning in

Diamond, A.D., and Hsu, J.T., 1990. Protein Partitioning in PEG/Dextran Aqueous Two-


I would like to sincerely thank my advisor Dr. Charles E. Glatz for his help, patience, guidance and support throughout the course of the work, especially during the difficult times.

I would like to thank the members of my program of study committee, Dr. John Kozak, Dr. Clark Ford, Dr. Peter Reilly, Dr. Kurt Hebert and Dr. Zifko Nikolov for their support and suggestions.

I would like to thank all my research group members who have helped me get started in the lab and whose suggestions and discussions were always valuable.

Finally, I would like to express my gratitude to my wife, for her trust, support, and encouragement through all these years and to my parents, who have always believed in me.