The UNIFAC-NRF activity coefficient model based on group contribution for partitioning of proteins in aqueous two phase (polymer + salt) systems

Ali Haghtalab [a,*], Babak Mokhtarani [a,1]

[a] *Department of Chemical Engineering, Tarbiat Modarres University, P.O. Box: 14115-175, Tehran, Iran*

Abstract

The group contribution model, UNIFAC-NRF, was applied for correlation of partition coefficient of proteins. This model was already developed for aqueous two phase of (polymer + salt) systems. The protein molecules were divided into some patches assumed to have interaction with the other species in aqueous (polymer + salt) systems. The binary interaction parameters were used for prediction of protein partitioning in aqueous two phase systems. These parameters were obtained by correlating the binary electrolyte solution and ternary aqueous two phase systems. The results of UNIFAC-NRF model are in a very good agreement with the experimental data for partitioning of lysozyme in both (PEG + K_2HPO_4 + water) and (PEG + Na_2SO_4 + water) systems at different pHs. The comparison of the results, which were obtained by both UNIFAC-NRF and the VERS models, shows that the present group contribution model can correlate the partitioning of protein in ATPS better than the VERS model.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Aqueous two phase system; Lysozyme; Partition coefficient; Activity coefficient; UNIFAC-NRF

1. Introduction

Aqueous two phase extraction systems are used for separation, concentrating and purification of proteins, cell organelles and other biological products. Aqueous two phase systems are particularly suited for separation and purification of proteins from crude material like cell extracts, fermentation broth, and culture filtrate. Beijerinck [1] first reported the liquid phase separation in such systems as early as in 1896. In 1956, Albertsson [2] based the separation and purification of a wide range of different biological materials such as proteins, viruses, DNA, etc. In recent years, many researchers investigated the separation of proteins and the other biomolecules. Kula *et al.* [3–5] have developed a number of protocols for protein extraction and purification in aqueous two phase systems. Huddleston and Lyddiatt [6–9] have investigated the recovery of proteins in these systems and studied the effect of parameters such as pH and polymer concentration on partitioning of proteins. Asenjo and co-workers investigated the partitioning of α-amylase, subtilisin, trypsin inhibitor and some enzymes in aqueous two phase systems of (PEG + phosphate). They have also investigated the effect of NaCl salt on partitioning in these systems [10–13]. Haghtalab *et al.* [19] investigated the partitioning of lysozyme, bovine serum albumin and α-amylase in aqueous two phase systems of (PEG + K_2HPO_4 + water) and (PEG + Na_2SO_4 + water) at $T = 298.15$ K and the effects of pH and salt concentration on protein partitioning were also studied.

Kaul *et al.* [14] studied the kinetic of phase separation for these systems. The application of these systems for the extraction of recombinant proteins was also developed [15,16]. Walter and Johansson [17] reviewed protein partitioning using aqueous two phase systems.

For the design of such extraction processes, a thermodynamic model for correlating/predicting (liquid + liquid) equilibria is of great importance. Such models can only be developed and tested if sufficient experimental data on well-defined model systems are available. However, the experimental data for protein partitioning are insufficient and the effect of important parameters such as pH on protein partitioning has not been investigated very well. To model partitioning of proteins, it is necessary to contribute the electrostatic interaction between molecules of proteins and other species in aqueous solution. The protein molecules behave as a charged species and the net charge of protein molecule has a significant effect on protein partitioning. The globular structure of proteins results in nonuniformity of the distribution of charges on the surface of protein molecules so, their thermodynamic behavior in aqueous solution is more complicated than are the other ionic fluids. In this work the UNIFAC-NRF model, which was proposed for aqueous two phase system of (PEG + salt + water) system, was extended for partitioning lysozyme [18]. The protein molecule was considered to consist of some patches, with charged or neutral species. The variation of protein net charge number was also studied at different pH values, and the number of charged groups was related to the net charge number of protein. Haghtalab *et al.* extended the VERS model for correlating protein partitioning in the aqueous two phase systems of (PEG + K_2HPO_4 + water) and (PEG + Na_2SO_4 + water) at $T = 298.15$ K with different pH and salt concentration[19]. Recently, Haghtalab and Mokhtarani [18] developed the UNIFAC-NRF model for correlation of phase partitioning in these aqueous two phase systems. In this work, this model was extended for protein partitioning in such aqueous two phase systems and the results of present moded compared with VERS model.

2. Thermodynamic framework

It was shown in previous work [18], that the activity coefficient of component "*i*" can be expressed as sum of two contributions:

$$\ln \gamma_i = \ln \gamma_i^{DH} + \ln \gamma_i^{UNIFAC-NRF}, \tag{1}$$

where the first term accounts for the contribution of long-range electrostatic interactions and the second term denotes the contribution of local composition based on the UNIFAC-NRF equation. The UNIFAC-NRF activity coefficient of component *i* can be written as

$$\ln \gamma_i^{UNIFAC-NRF} = \ln \gamma_i^C + \ln \gamma_i^R, \tag{2}$$

where the terms on the right-hand side are the combinatorial and residual parts of UNIFAC-NRF model. The combinatorial activity coefficient is presented as

$$\ln \gamma_i^C = \ln \left(\frac{\phi_i}{x_i} \right) + \frac{zq_i}{2} \ln \frac{\theta_i}{\phi_i} + l_i - \left(\frac{\phi_i}{x_i} \right) \sum_{j=1}^{n} (x_j l_j), \tag{3}$$

where ϕ_i and θ_i are the volume and area fractions of molecule *i*, respectively. The residual part of UNIFAC-NRF can be written as

$$\ln \gamma_i^R = \sum v_k^{(i)} \cdot \left[\ln \Gamma_k - \ln \Gamma_k^{(i)} \right], \tag{4}$$

where $v_k^{(i)}$ is the number of group *k* in molecule *i*, Γ_k and $\Gamma_k^{(i)}$ are the activity coefficient of group *k* in the mixture and solution of pure molecules *i*, respectively. The activity coefficient of the group using UNIFAC-NRF was presented as [18]

$$\ln \Gamma_k = Q_k \cdot \left[1 + \ln \xi_{kk} - \sum_{j=1} \Theta_j \xi_{kj} + (1 - \Theta_k) \cdot \sum_{j \neq k} \Theta_j \cdot \right.$$
$$\left. \ln \frac{\xi_{kj} \cdot \xi_{jk}}{\xi_{kk} \cdot \xi_{jj}} - \frac{1}{2} \sum_{l \neq k} \sum_{\substack{m \neq k \\ m \neq l}} \Theta_l \cdot \Theta_m \cdot \ln \frac{\xi_{ml} \cdot \xi_{lm}}{\xi_{mm} \cdot \xi_{ll}} \right], \tag{5}$$

$$\xi_{kk} = 1 / \sum_{\substack{k \neq j \\ \text{for (ion)}}} \theta_j \psi_{jk}, \tag{6}$$

$$\Psi_{ij} = \exp(-a_{ij}/T), \tag{7}$$

$$\xi_{ij} = \psi_{ij} \xi_{jj}, \tag{8}$$

where a_{ij} is the adjustable interaction parameter, Θ_{ij} and ξ_{ij} are the local area fraction and non-random factor of group *i*, surrounding the central group *j*, respectively. The Θ_i is the surface area fraction of group *i*, Q_k is relative surface area parameter and X_k is the mole fraction of group *k*.

For the contribution of long-range interactions to the excess Gibbs free energy, the unsymmetrical normalization is used; it is necessary to normalize the contribution due to short-range interactions on the same basis. Thus, we may use

$$\ln \gamma_i^* = \ln \gamma_i - \ln \gamma_i^\infty, \tag{9}$$

where γ^∞ is the activity coefficient at infinite dilution and can be derived from the boundary conditions at $x_s \to 0$, $x_p \to 0$, $x_w \to 1$ as the following:

$$\ln \gamma_i^\infty = \ln \frac{r_i}{r_w} + 5q_k \times \ln \frac{q_i r_w}{r_i q_w} + l_i - \frac{r_i}{r_w} \cdot l_w +$$
$$\sum_k v_k^{(i)} Q_k \left[(1 + \ln \psi_{kw} - \psi_{kw}) - \ln \Gamma_k^{(i)} \right]. \tag{10}$$