Research note

New Perturbation Model for Prediction of Amino Acid and Peptide Activity Coefficients

M. R. Dehghani[1*], E. Salehi[1,2], A. R. Fazlali[2]

1- Thermodynamics Research Laboratory, School of Chemical Engineering, Iran University
of Science and Technology, Tehran, Iran
2- Department of Chemical Engineering, Faculty of Engineering, Arak University, Arak, Iran

Abstract

In this work, a new thermodynamic model based on the perturbation theory is presented. A new hard spheres equation of state as a reference term is applied to correlate the activity coefficient of amino acids and peptides in binary aqueous solutions. The new hard sphere equation of state has been recently proposed by Dehghani and Modarress [11] and has been applied for different theories and showed excellent capability. In this model dipole-dipole and Lennard-Jones interactions are considered. The results have been compared with similar models and it is shown that application of the new hard spheres equation of state has caused an improvement in the results of perturbation model.

Keywords: *Activity Coefficient, Hard Sphere, Perturbation Model, Amino Acid*

1. Introduction

In recent years, production of biochemicals and pharmaceutical materials has significantly increased. The high cost of separation and concentration of biomolecules from dilute aqueous solution (which can be as high as 90% of their total manufacturing cost [1]) has encouraged researchers to find new methods for purification of these solutions. Fractional precipitation and crystallization methods are often used for separation and concentration of the biomolecules [1]. In this regard, thermodynamic studies such as experimental

and theoretical studies for correlation and prediction of activity coefficient of amino acids as the simplest biomolecules are valuable.

Thermo physical properties are essential factors in the design of the separation processes such as crystallization, distillation and liquid-liquid extraction. During the past decades, many scientists have focused on determination of activity coefficients of amino acids as well as peptides in aqueous solutions by experimental and theoretical studies [2-12]. In 1989, local composition model (NRTL) was used by Chen et al. [2]

∗ Corresponding author: m_dehghani@iust.ac.ir

for prediction of phase partitioning and amino acid solubility. After that, conventional thermodynamic models including group contribution methods, Wilson, UNIQUAC and UNIFAC models were used by Nass [6], Peres et al. [7], Jin et al. [8] and Gupta et al. [9], Pinho et al. [10], Kuramochi et al. [11] respectively. Recently, the perturbation theory has also been incorporated into the amino acid systems by Khoshkbarchi and Vera [12] to correlate the activity coefficients of amino acids and peptides in the aqueous solutions.

In this work, we present a new two parametric model based on perturbation theory for prediction of amino acids and peptides activity coefficients. In this model a new hard sphere equation of state is used as a reference system. This hard sphere equation of state was recently proposed by Dehghani and Modarress [13] and has shown significant ability in comparison with other hard sphere equation of states.

2. Theory of the model
The perturbation theory was initially developed by Zwanzig [14] and was extended to liquids by Barker and Henderson [15]. In this theory, the total interaction energy of the system is divided into two reference and perturbation contributions. The perturbation theory has the advantage of incorporating various types of interactions easily. In our model different forms of interactions have been considered. In this work a primitive perturbation model has been utilized to ignore the hydrogen bonding between solvent and solute. However, the amino acids, and especially peptides, are the

chain like molecules; nevertheless, in this model they have been considered as hard spheres containing dipole moments just for the simplicity of the model. Based on these assumptions, the residual chemical potential of the amino acids can be defined as follows:

$$\frac{\mu_i^r}{kT} = \left(\frac{\mu_i^r}{kT}\right)^{REF} + \left(\frac{\mu_i^r}{kT}\right)^{PER} \quad (1)$$

Where the superscripts *REF.* and *PER.* denote the contributions of the reference and perturbation term, respectively. k is Boltzmann constant ($1.38 \times 10\text{-}23 \ JK^{-1}$) and T is the absolute temperature. Generally hard sphere equations of state are used as reference term. There are several hard sphere equations of state in the literature, in this paper we have focused on using a new hard spheres equation of state which was recently presented by Dehghani and Modarress [13]. We have utilized this Hard sphere EOS because of its unique abilities. This hard sphere equation of state was built based on new highly optimized molecular dynamics simulation data. It is capable of meeting both the low density and closed packed limits of compressibility factor [13]. In this model, the compressibility factor for pure hard spheres is presented as follows:

$$Z_p = [1 - 1.038\zeta - 0.364\zeta^2 - 0.706\zeta^3 + 0.718\zeta^4 + 0.686\zeta^5 + 0.352\zeta^6 - 1.378\zeta^7 + 1.911\zeta^9 - 1.299\zeta^{11}]/(1-\zeta)^4 \quad (2)$$

Where ζ is equal to v_0/v (v_0 is the occupied volume by hard spheres and v is the total volume). The mathematical equation for ζ is defined as Eq. (3):

$$\zeta = \frac{1}{\sqrt{2}} \rho \sigma^3 \qquad (3)$$

In the above equation σ and ρ are the size parameter and number density respectively. Applying pure hard sphere compressibility factor, the residual Helmholtz free energy A^r, and the chemical potential, μ^r, can be determined as Eq. (5) and (6) respectively.

$$\frac{A^{r,HS}[T,V,N]}{NkT} = \int_0^\rho \left[\frac{Z_p - 1}{\rho}\right] d\rho \qquad (4)$$

The result of a combination of Eq. (2) and (4) has been presented as follows [13]:

$$\frac{A^{r,HS}[T,V,N]}{NkT} = -0.186\zeta^7 - 0.866\zeta^6$$
$$-2.216\zeta^5 - 4.58\zeta^4 - 9.245\zeta^3 - 19.845\zeta^2$$
$$-53.93\zeta - \frac{0.359}{(1-\zeta)^3} + \frac{1.045}{(1-\zeta)^2} - \frac{15.539}{(1-\zeta)}$$
$$-70.46\ln(1-\zeta) + 14.533 \qquad (5)$$

$$\frac{\mu_i^{r,HS}}{kT} = \frac{A^{r,HS}}{NkT} + \rho\left(\frac{\partial(A^{r,HS}/NkT)}{\partial\rho_i}\right)_{T,V,\rho_{j\neq i}} \qquad (6)$$

The perturbation term of chemical potential in our model has been defined using an equation based on Barker-Henderson theory [15]. According to Tiepel-Gubbins [16] the chemical potential term can be calculated utilizing Eq. (7):

$$\mu_i^{Per} = 4\pi\rho \int_\sigma^\infty u(r)g^{HS}(r)r^2 dr \qquad (7)$$

In which; r, $u(r)$, and $g^{HS}(r)$ are intermolecular distance, intermolecular

potential and radial distribution function, respectively. Different types of intermolecular interactions can be considered. In this work the most important ones such as dipole-dipole and dispersion interaction energies have been considered. Lenard-Jones model has been employed for considering dispersion energy (Eq. (8)). Dipole-dipole interactions have been defined as Eq. (9) [17].

$$u^{L-J}(r) = 4\varepsilon\left(\frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6}\right) \qquad (8)$$

$$u^{D-D} = -\frac{D^4}{3(4\pi\varepsilon_0\varepsilon_r)^2 kTr^6} \qquad (9)$$

Where, ε, D, ε_0 and ε_r are depth potential well, dipole moment, permittivity of vacuum and relative dielectric constant, respectively. Combining Eq. (8) and (9) the perturbation part in residual chemical potential will be derived as Eq. (10).

$$\mu_i^{r,Per} = -4\pi\rho\left[\frac{8}{9}\varepsilon\sigma^3 + \frac{D^4}{9(4\pi\varepsilon_0\varepsilon_r)^2 kT\sigma^6}\right] \qquad (10)$$

In the above equation ρ is number density. Unsymmetrical activity coefficient of solution, γ_i, is related to the residual chemical potential according to Eq. (11):

$$\ln\gamma_i = \frac{\mu_i - \mu_i^{id}}{kT} \qquad (11)$$

In which μ_i^{id} is chemical potential of component i in ideal solution. Residual chemical potential of amino acid is presented as follows:

$$\frac{\mu_A^r}{kT} = \ln \gamma_A = \ln \gamma_A^{HS} + \ln \gamma_A^{PER} \qquad (12)$$

Solution density is a necessary physical property data in calculation of amino acid and peptide activity coefficients and its accuracy is important. Although in previous models pure solvent density was applied as solution density, in this work Eq. (13-14) proposed by Mishra-Ahluwalia [18] was employed to predict solution density.

$$d = \frac{1000 + m.M}{m.\varphi + 1000/d_0} \qquad (13)$$

$$\varphi = \overline{V}^0 + S_v.m \qquad (14)$$

Where m and M are the molality and molecular weight of amino acid, respectively, and d_0 is density of pure water at experimental condition.

3. Parameter estimation

In the presented model there are two adjustable parameters, size and depth of potential well parameters, which should be determined based on available experimental data on activity coefficient. Dipole moments of amino acids and peptides are constant values. Dipole moments cannot be measured experimentally [17] and must be calculated through quantum mechanical approach. Usually, modeling softwares such as Hyperchem molecular modeling can be used for this purpose [12].

The estimation of parameters is a very important procedure, because its results will affect the predictability of the model. In order

to avoid the local optimization instead of global optimization, Davidon-Fletcher-Powell (DFP) [13] algorithm was used to avoid the method of direct inversion of Hessian matrix. DFP is a well known procedure which approximates the inverse Hessian matrix to find the global minimum. Adjustable parameters were determined through minimization of objective function. Objective function was defined as follows:

$$OF = \sum_{i=1}^{n} \frac{\left(\gamma_i^{\exp} - \gamma_i^{cal} \right)^2}{n} \qquad (15)$$

Available experimental data on activity coefficient is molality based while calculated activity coefficients are based on mole fraction, in this regard Eq. (16) is used to change the experimental activity coefficient from molality based to mole fraction based.

$$\ln \gamma^{(x)} = \ln \gamma^{(m)} + \ln(1 + 0.001 M_s \sum_i m_i) \qquad (16)$$

In which M_s is molecular weight of solvent. The superscripts x and m denote the molality and the mole fraction base activity coefficients, respectively.

Finally, deviation of the presented model from the experimental data was calculated using root mean square deviation (RMSD) as shown below:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n} (\gamma^{\exp} - \gamma^{cal})}{n}} \qquad (17)$$

Where γ^{exp} is the experimental value of activity coefficient, γ^{cal} is the calculated

activity coefficient by Eq. (12), and n is number of data.

4. Results and discussion

In Fig.1 and 2, the activity coefficient of different amino acids such as amino butyric acid, alanine, hydroxiporline and threonine have been correlated using a new model. It can be seen that the presented model could correlate the experimental values efficiently. In Fig. 3 the same results have been presented for peptides. As it is shown, the overall fitting is quite good and from the results it can be concluded that the applied modifications have improved the fitting procedure. In Tables 1 and 2, calculated values for size parameter and depth potential well have been presented. All adjusted values are in the acceptable range. This subject can be referred to the applied physically meaningful assumptions in this model. The obtained results have been compared with other models such as those presented by

Khoshkbarchi and Vera [12], Mortazavimanesh et. al.[19], Chen et al.[5], Gupta et al.[9] and finally Pinho[10]. Calculated RMSD indicates that this model can correlate experimental data more accurately than the other ones. Table 2 shows that the presented model could correlate the activity coefficient of peptides accurately. In the case of amino acids, the average RMSD in the present work is 0.34 while for other models such as Khoshkbarchi, Mortazavimanesh, Chen, Gupta, Heidemann and Pinho it is 0.89, 0.6, 1.51, 7.85 and 8 respectively. Similar results have been obtained for peptides. It is worth mentioning that in this work just two adjustable parameters have been utilized for correlation. Meanwhile, it can be concluded that applying new hard sphere equation of state as well as modification on solution density (as other researchers used density of pure water) improved the results of perturbation theory in comparison with other similar models.
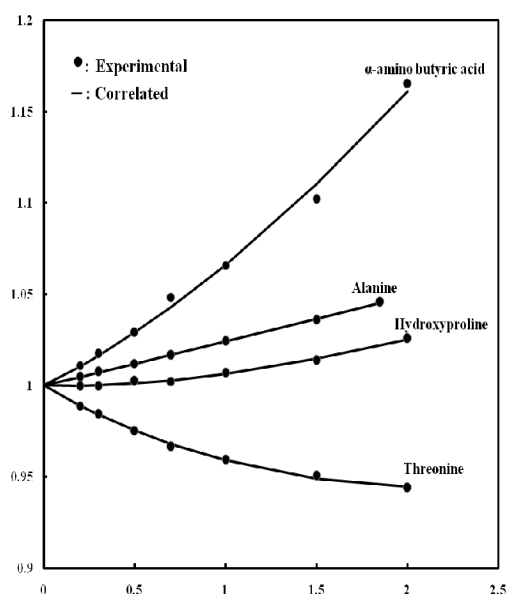

Figure 1. Activity coefficient of amino acids versus their molalities
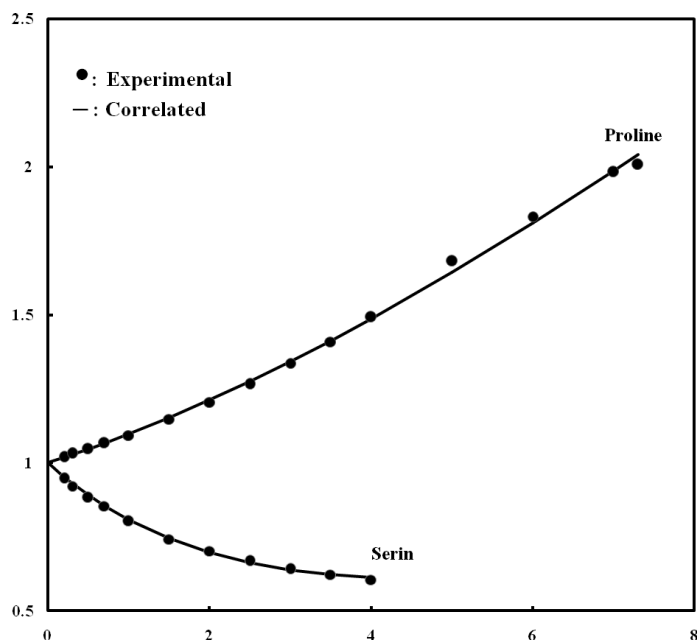(all experimental data used were output from Fasman [20])

Figure 2. Activity coefficient of amino acids versus their molalities
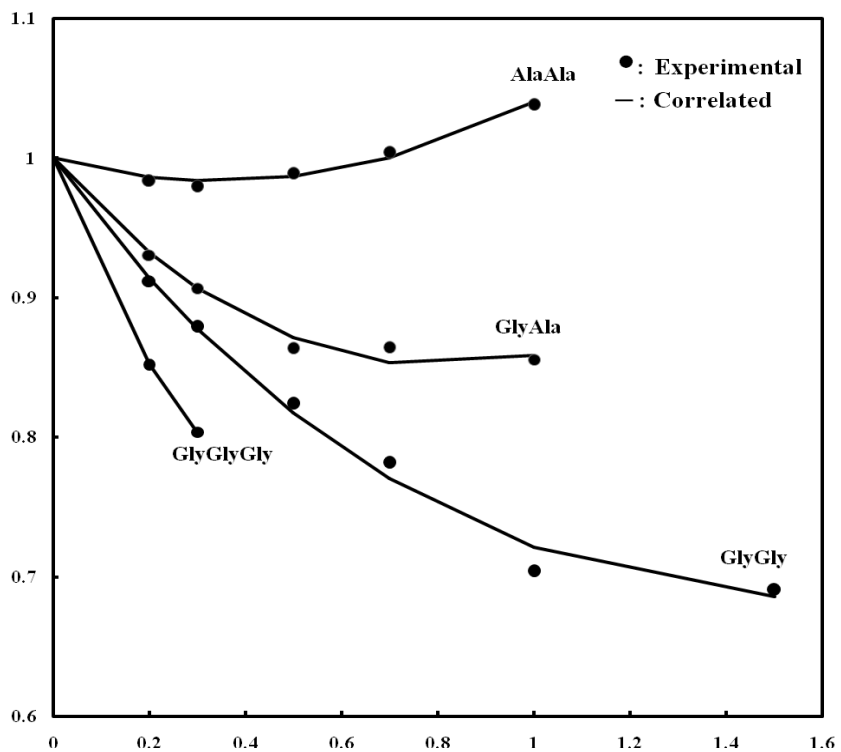(all experimental data used were output from Fasman [20])

Figure 3. Activity coefficient of peptides versus their molalities,
(all experimental data used were output from Fasman [20])

Table 1. Calculated σ, ε/k and RMSD from Correlation of Amino Acids experimental Activity Coefficients

Rmsd × 100[6]	Rmsd × 100[5]	Rmsd × 100[4]	Rmsd × 100[3]	Rmsd × 100[2]	Rmsd × 100[1]	$\sigma \times 10^{10}$ (m)	ε/k(K)	Amino acid
0.19	8.97	0.04	0.03	0.33	0.05	3.33	60.77	Alanine
17.84	17.34	0.32	-	0.37	o.4	4.8	87.06	α-Amino n-butyric acid
1.67	4.20	2.07	0.66	0.86	0.39	4.71	174.92	Glycine
0.39	0.06	0.36	-	0.09	0.07	4.24	107.57	Hydroxyproline
1.21	3.01	1.30	1.38	4.12	1.55	4.5	62.73	Proline
0.24	3.32	2.80	0.24	1.16	0.54	5.03	196.74	Serine
26.12	13.78	0.70	0.25	0.15	0.08	4.54	135.21	Threonine
8	7.85	1.51	0.60	0.89	0.38			Average rmsd

(1) Calculated in this work; (2) Khoshkbarchi-Vera [12]; (3) Mortazavi manesh et al. [19]; (4) Chen et al. [5]; (5) Gupta and Heidemann [9]; (6) Pinho [10]

Table 2. Calculated σ , ε/k and RMSD from Correlation of Peptides Experimental Activity Coefficients.

Rmsd × 100[5]	Rmsd × 100[4]	Rmsd × 100[3]	Rmsd × 100[2]	Rmsd × 100[1]	$\sigma \times 10^{10}$ (m)	ε/k(K)	Peptide
0.52	2.04	0.34	0.43	0.32	6.66	117.85	Alanylalanine
28.22	2.92	0.41	0.60	0.39	6.91	149.58	Alanylglycine
0.59	2.44	-	0.57	0.61	7.22	153.22	Glycylalanine
17.06	3.25	0.99	1.22	0.91	6.58	174.28	Glycylglycine
10.00	0.67	0.00	0.00	0.00	8.56	160.25	Tri-glycine
11.29	2.26	0.50	0.56	0.44			Average rmsd

(1) Calculated in this work; (2) Khoshkbarchi-Vera [12]; (3) Mortazavi manesh et al. [19]; (4) Chen et al. [5]; (6) Pinho [10]

5. Conclusions

In this work a new model based on perturbation theory was developed for correlation of experimental activity coefficients of amino acids and peptides. In our model a new hard sphere equation of state was used as a reference system. Dipole-dipole and dispersion interactions were considered as perturbation terms. In order to improve the accuracy of the model an empirical equation was used for prediction of solution density. The results of this work were compared with previous works and it was shown that the presented model was successful in correlation of the experimental data. This subject can be referred to the following items: first of all, application of new highly optimized molecular dynamics

simulation data in constructing the new hard sphere equation of state (HS-EOS). Secondly, capability of the applied HS-EOS, to meet low density as well as closed packed limits of compressibility factor. Results of the model denote that the effect of reference term and solution density in chemical potential of amino acids is significant. Compared to similar models, the modifications caused an improvement in the results of the perturbation model.

Nomenclature

A^r	Residual chemical potential
d	Solution density
d_0	Density of pure water
D	Dipole moment
$g^{HS}(r)$	Radial distribution function
k	Boltzmann constant
m	Molality of amino acid
M	Molecular weight of amino acid
M_S	Molecular weight of solvent
n	Number of data
N	Number of spheres
OF	Objective function
r	Intermolecular distance radial distribution function
T	Absolute temperature
u(r)	Intermolecular potential
V	Molar volume
γ	Activity coefficient
ε	Depth of potential well
ε_0	Permittivity of vacuum
ε_r	Relative dielectric constant
μ_i^r	Residual chemical potential
ρ	Number density
σ	Size parameter
Z_P	Compressibility factor

References

[1] Belter, P. A., Cussler, E. L. and Hu, W. S., Biosepartion, John Wiley, New York, (1998).

[2] Dehghani, M. R., Modarress, H. and Bakhshi, A., "Modeling and prediction of activity coefficient ratio of electrolytes in aqueous electrolyte solution containing amino acids using artificial neural network", Fluid Phase Equilib, 244, 153-159, (2006).

[3] Dehghani, M. R., Modarress, H. and Monirifar, M., "Measurement and modeling of mean activity coefficients of NaBr and amino acid in (Sodium Bromide+Potassium Phosphate + Glycine + water) at 298.15 and 308.15 K", J. Chem. Thermodynamics, 37, 1305-1311, (2005).

[4] Dehghani, M. R., Modarress, H. and Monirifar, M., "Measurement and modelling of mean activity coefficients of aqueous mixed electrolyte solution containing glycine", J. Chem. Thermodynamics, 38, 1049-1053, (2006).

[5] Chen, C. C, Zhu, Y. and Evans, L. B., "Phase partitioning of biomolecules, solubilities of amino acids", Biotechnol. Prog., 5, 111, (1989).

[6] Nass, K. K., "Representation of the solubility behavior of amino acids in water", AIChE J., 34, 1257-1266, (1988).

[7] Peres, A. M. and Macedo, E. A., "Representation of solubilities of amino-acids using the uniquac modelfor electrolytes", Chem. Eng. Sci., 49,

3803-3812, (1994).

[8] Jin-Chen, L., Jiu-Fang, L. and Yi-Gui, L., "Study on the activity coefficients and the solubilities of amino acids in water by the perturbation theory", Fluid Phase Equilib., 142, 67-82, (1998).

[9] Gupta, R. B. and Heidemann, R. A., "Solubility models for amino acids and antibiotics", AIChE J., 36, 333, (1990).

[10] Pinho, S. P., Silva, C. M. and Macedo, E. A., "Solubility of amino acids: A group contribution model involving phase and chemical equilibria", Ind. Eng. Chem. Res., 33, 1341-1347, (1994).

[11] Kuramochi, H., Noritomi, H., Hoshino, D. and Nagahama, K., "Measurements of solubilities of two amino acids in water and prediction by the UNIFAC model", Biotechnol. Prog. 12, 371-379, (1996).

[12] Khoshkbarchi, M. K. and Vera, J. H., "A simplified perturbed hard sphere model for the activity coefficients of amino acids and peptides in aqueous solutions", Ind. Eng. Chem. Res., 35, 4319-4327, (1996).

[13] Dehghani, M. R. and Modarress, H., "Modified equation of state for pure and hard sphere mixtures in mean ionic activity coefficient calculations by mean spherical approximation model", J. Molec. Liquids, 142, 45-52, (2008).

[14] Zwanzig, R.W., "High-temperature equation of state by a perturbation method. I. nonpolar gases", J. Chem. Phys., 22, 1420, (1954).

[15] Barker, J. A. and Henderson, D., "Perturbation theory and equation of state for fluids. II. A successful theory of fluids", J. Chem. Phys., 47, 4714, (1967).

[16] Tiepel, T. M. and Gubbins, K. E. "Thermodynamic properties of gases dissolved in electrolyte solutions", Ind. Eng. Chem. Fundam., 12, 18, (1973).

[17] Maitland, G. C., Rigby, M., Smith, E. B. and Wakeham. W. A., Intermolecular forces, Clarendon Press, Oxford, (1981).

[18] Mishraw, A. K. and Ahluwalia, J. C., "Apparent molal volumes of amino acids, N-acetylamino acids, and peptides in aqueous solutions", J. Phys. Chem., 8886-92, (1984).

[19] Mortazavi-Manesh, S., Ghotbi, C. and Taghikhani, V., "A new model for predicting activity coefficients in aqueous solutions of amino acids and peptides", J. Chem. Thermodynamics, 35, 101–112, (2003).

[20] Fasman, G. D., Handbook of bio-chemistry and molecular biology, 3rd ed., Physical and Chemical Data, CRC Press, Cleveland, Vol. 1, (1976).