

A new approach for the thermodynamic modeling of the solubility of amino acids and β -lactam compounds as a function of pH



Luís Fernando Mercier Franco^a, Silvana Mattedi^b, Pedro de Alcântara Pessôa Filho^{a,*}

^a Bioprocess Engineering Group, Chemical Engineering Department, School of Engineering, Universidade de São Paulo, Caixa Postal 61548, 05424-970 São Paulo, SP, Brazil

^b Chemical Engineering Graduate Program, Polytechnic School, Universidade Federal da Bahia, Brazil

ARTICLE INFO

Article history:

Received 24 January 2013

Received in revised form 5 June 2013

Accepted 6 June 2013

Available online 15 June 2013

Keywords:

Solid–liquid equilibrium

Modeling

Amino acids

β -Lactam compounds

ABSTRACT

A new model to describe the solubility of amino acids and β -lactam compounds as a function of pH is proposed. The description of the solid–liquid equilibrium entails simultaneous calculation of phase and chemical equilibria. The liquid-phase non-ideality was accounted for through an extended Pitzer model for polyelectrolyte solutions. A statistical thermodynamic interpretation of Pitzer's binary interaction parameters is also presented. The model was successfully applied to different systems containing amino-acids, as well as glycine and its oligopeptides and β -lactam compounds such as ampicillin and 6-aminopenicillanic acid.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

When modeling the solubility of amino acids or β -lactam antibiotics as a function of pH, the most important aspect is the liquid phase modeling, since solid phase can be assumed to be unique and constituted only by solute molecules. A specific issue concerning the solubility of amino acids is the fact that in aqueous solutions they dissociate, so that a single amino acid can be found in three different states: as a cation, as an anion and as a zwitterion – the neutral nonpolar species is virtually non-existent in the solution. Many strategies have been proposed to calculate amino acid activity coefficients in liquid phase: through the Poisson–Boltzmann equation [1], through excess Gibbs energy models [2–8], through group-contribution models [9,10], through statistical thermodynamic approaches [11] and through volumetric equations of state [12–14].

To the best of our knowledge, the first model to describe amino acid solubility curve as a function of pH was based entirely upon considerations about chemical dissociation [15] without considering liquid-phase non-ideality. The solubility curve of tyrosine in aqueous solution as a function of pH was calculated with good agreement with experimental data.

The work by Kirkwood [1] is among the first ones to describe solutions containing zwitterions. It is based on the analytic

solution of the linearized Poisson–Boltzmann equation, considering the three directions on the Laplacian operator for a spherical zwitterion. The author obtained an expression that relates the logarithm of an ion activity coefficient to a series involving multipole moments. The first term of this series is analogous to the Debye–Hückel term, being null for zwitterions, as they have no net charge. With some simplifications, the author also showed the application of this approach to an ellipsoid zwitterion [16].

Nass [2] modeled the solubility of amino acids as a function of pH assuming that the excess Gibbs energy is given by the sum of chemical and physical contributions. The chemical contribution was calculated considering amino acid dissociation, mass balance equations and phase equilibrium conditions, whereas the physical contribution was calculated through Wilson's equation [17] for the binary mixture of water and amino acid. The experimental data were well correlated, at the expense of a high number of adjustable parameters. Chen et al. [3] considered that the excess Gibbs energy could be split in three terms, two of them related to long-range interactions, viz., a Pitzer–Debye–Hückel and a Born terms, and another one to account for short-range interactions, the NRTL equation [18]. For the activity coefficient of the zwitterions, the Pitzer–Debye–Hückel contribution becomes null. The authors showed that the experimental solubility data were satisfactorily correlated with only one or two adjustable parameters for each compound.

The work by Gupta and Heidemann [9] laid the foundation for most ensuing works. Their approach relates chemical dissociation of amino acids and the solid–liquid equilibrium condition. The

* Corresponding author. Tel.: +55 11 3091 1106; fax: +55 11 3091 2284.

E-mail address: pedro.pessoa@poli.usp.br (P.d.A. Pessôa Filho).

List of symbols*Latin letters*

a	universal parameter of the Pitzer's model ($\text{kg}^{1/2} \text{mol}^{-1/2}$)
a_i	activity of component i
b	universal parameter of the Pitzer's model ($\text{kg}^{1/2} \text{mol}^{-1/2}$)
B_{ij}	osmotic second virial coefficient for interaction between components i and j (m^3)
$C_{ij}^{(1)}$	adjustable parameter for the Pitzer's model (kg mol^{-1})
$C_{ij}^{(2)}$	adjustable parameter for the Pitzer's model ($\text{kg mol}^{-1} \text{K}^{-1}$)
d_s	solvent density (kg m^{-3})
e	electric charge (As)
I	ionic strength (mol kg^{-1})
k_B	Boltzmann's constant ($\text{m}^2 \text{kg s}^{-2} \text{K}^{-1}$)
K_{A1}	equilibrium constant of acid protonation
K_{A2}	equilibrium constant of amine protonation
K_D	equilibrium constant of neutral molecules
m_0	molality of electrically neutral amino acids (mol kg^{-1})
m_i	molality of component i (mol kg^{-1})
m^*	molality of component i in the reference state (mol kg^{-1})
M	number of experimental data sets
M_s	solvent molecular mass (kg mol^{-1})
N	number of experimental data points
N_A	Avogadro's number (mol^{-1})
r	radial coordinate (m)
R	gas constant ($\text{kg m}^2 \text{mol}^{-1} \text{K}^{-1} \text{s}^{-2}$)
S	solubility (mol kg^{-1})
S_i^{exp}	experimental solubility (mol kg^{-1})
S_i^{calc}	calculated solubility (mol kg^{-1})
T	temperature (K)
W_{ij}	potential of mean force between i and j ($\text{kg m}^2 \text{mol}^{-1} \text{s}^{-2}$)
z_i	electric charge of component i
z_j	electric charge of component j

Greek letters

α_{AA-}	ionized fraction of amino acid in anionic form
α_{AA+}	ionized fraction of amino acid in cationic form
γ_0	electrically neutral amino acid activity coefficient
γ_{AA-}	zwitterionic activity coefficient
γ_i	activity coefficient of component i
ε	van der Waals attractive interaction parameter ($\text{m}^2 \text{kg s}^{-2}$)
ε_0	vacuum permittivity ($\text{A}^2 \text{s}^4 \text{kg}^{-1} \text{m}^{-3}$)
ε_r	dielectric constant
κ	inverse of Debye's length
λ_{ij}	Pitzer's binary interaction parameter (kg mol^{-1})
Λ_{ijk}	Pitzer's ternary interaction parameter ($\text{kg}^2 \text{mol}^{-2}$)
μ_i^L	chemical potential of component i in the liquid phase ($\text{kg m}^2 \text{mol}^{-1} \text{s}^{-2}$)
μ_i^S	chemical potential of component i in the solid phase ($\text{kg m}^2 \text{mol}^{-1} \text{s}^{-2}$)
μ_i^*	chemical potential of component i in the reference state ($\text{kg m}^2 \text{mol}^{-1} \text{s}^{-2}$)
ρ_i	number density of component i (m^{-3})
ρ_j	number density of component j (m^{-3})
σ	van der Waals diameter (m)
ϕ_0	fraction of electrically neutral amino acids

ϕ_{AA}	fraction of nonpolar amino acid molecules
ϕ_{AA-}	fraction of anionic amino acid molecules
ϕ_{AA+}	fraction of cationic amino acid molecules
ϕ_{AA-}	fraction of zwitterionic amino acid molecules

non-ideality of aqueous solutions of native and modified amino acids was calculated through the modified UNIFAC model of Larsen et al. [19] with new groups such as glycine and proline. Following their work, Pinho et al. [10] proposed a UNIFAC model combined with a Debye–Hückel term to calculate the activity coefficients of amino acids; the solubility of glycine was modeled as a function of pH with very good agreement with experimental data. Nevertheless, some authors question the introduction of the Debye–Hückel term in this case [5,11]. Rodríguez-Raposo et al. [4] fitted osmotic coefficients of glycine in aqueous solution using the Pitzer equation and calculated the activity coefficient of glycine in aqueous solution, with very good agreement to the experimental data.

Khoshkbarchi and Vera [5] split the excess Gibbs energy in long-range and short-range interaction contributions. They tested the models by Bromley [20] and their own [21] for the long-range interaction contribution and, for the short-range term, the NRTL and Wilson models [17]. Whichever the model chosen, the approach required only two parameters, and binary water–amino acid systems and ternary water–electrolyte–amino acid systems were well correlated.

Even though the basic equation to predict the glycine solubility as a function of pH used by Khoshkbarchi and Vera [11] follows the approach by Gupta and Heidemann [9], they introduce the use of the perturbation theory to calculate the non-ideality of the amino acid in the aqueous solution. They modeled amino acids as Lennard–Jones spheres with dipole–dipole interactions. Santana et al. [22] showed that this model fails to quantitatively predict the solubility of ampicillin, D-phenylglycine and 6-aminopenicillanic acid as a function of pH even adjusting five parameters. Pradhan and Vera [6], using an equation similar to that proposed by Gupta and Heidemann [9] with the NRTL equation, correlated the DL-alanine solubility as a function of pH, with good agreement with the experimental data in the vicinity of the isoelectric point.

Park et al. [12] correlated the activity coefficient of valine, α -amino butyric acid, alanine, glycine and glycylglycine in aqueous solution by applying the hydrogen-bonding lattice fluid equation of state, with no consideration about amino acid dissociation.

Xu et al. [7] proposed a modification of the Wilson model to calculate the excess Gibbs energy of amino acid solutions. They assumed that all amino acid molecules were in zwitterionic form in pure water, which somehow restricts the application of their model for calculating amino acid solubility at pH in the vicinity of isoelectric point.

Using the PC-SAFT equation of state and accounting for amino acid dissociation in solution, Fuchs et al. [13] described the solubility curve as a function of pH for glycine, DL-alanine and DL-methionine in aqueous solution with good agreement with experimental data. Similarly, Seyfi et al. [14], following Gupta and Heidemann [9] and applying SAFT equation of state, correlated the DL-methionine solubility curve as a function of pH and temperature with high accuracy.

Tseng et al. [8] investigated the solubility of DL-alanine, L-leucine, L-isoleucine, L-serine and DL-phenylalanine as a function of pH in aqueous solution. Considering amino acid dissociation in aqueous solutions and using the NRTL model, they obtained satisfactory results with two adjustable parameters. Dalrup et al. [23] modeled systems containing two amino acids in solution using PC-SAFT equation of state.

Thus, while models differ considerably on which contributions are considered in the activity coefficient model, the calculation of acid–base equilibrium remains as the fundamental step in describing the solubility curve.

2. Theoretical framework

The basis of the model is a recent work on modeling protein solubility curve as a function of pH [24]. In this case, the equation for amino acids (regardless of the residue) is just a particular case of the same equation with two or three ionizable groups.

The condition of solid–liquid equilibrium for a given system, in which solid phase is considered to be unique and to contain only molecules of solute, is:

$$\mu_i^S = \mu_i^L \quad (1)$$

wherein μ_i^S is the chemical potential of solute i in solid phase and μ_i^L is the chemical potential of solute i in liquid phase. Neglecting the effect of pressure on chemical potential and considering an isothermal process, the chemical potential of solid phase will not change, as its composition is fixed. Thus, the chemical potential of solute in liquid phase must remain constant, and the pH value does not alter the solute chemical potential in liquid phase.

Using molality scale [25]:

$$\mu_i^L = \mu_i^* + RT \ln a_i = \mu_i^* + RT \ln \left(\frac{\gamma_i m_i}{m^*} \right) \quad (2)$$

wherein μ_i^* is the standard state chemical potential of solute i in a hypothetical ideal solution with unit concentration ($m^* = 1.0 \text{ mol kg}^{-1}$), γ_i is the activity coefficient of solute i and m_i is the molality of i in liquid phase.

Amino acid or β -lactam antibiotics may have four different ionization states in aqueous solutions. Eq. (1) is valid only for those which are electrically neutral species. This restriction is due to the fact that solid phase is electrically neutral, and hence the phase equilibrium is established only by electrically neutral species of amino acids. Thus, the activity of electrically neutral molecules remains constant with any variation of pH. The experimentally determined solubility, however, is obtained considering the concentration of all kinds of amino acids molecules irrespective to their ionization state. Thus, the molality of the electrically neutral molecules is a fraction of the solubility:

$$m_0(\text{pH}) = \phi_0(\text{pH})S(\text{pH}) \quad (3)$$

wherein m_0 is the molality of the electrically neutral molecules of amino acids, ϕ_0 is the fraction of electrically neutral molecules of amino acids and S is the amino acid solubility. As the activity of electrically neutral molecules of amino acids is constant, its value at any pH is equal to its value at isoelectric point (pI). Thus, one can write:

$$S(\text{pH}) = \frac{\gamma_0(pI)\phi_0(pI)S(pI)}{\gamma_0(\text{pH})\phi_0(\text{pH})} \quad (4)$$

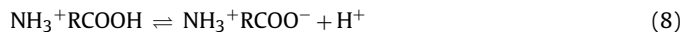
If the amino acid concentration in liquid phase is small, Henry's law applies and Eq. (4) can be rewritten simply as [24]:

$$S(\text{pH}) = \frac{\phi_0(pI)S(pI)}{\phi_0(\text{pH})} \quad (5)$$

2.1. Chemical equilibrium

Let us firstly consider amino acids with only two ionizable groups in its chain – extension for amino acids which have three ionizable groups is straightforward.

In aqueous solutions, the following equilibria take place:



There are four different species of amino acid molecules: the neutral non-ionized one (AA), the zwitterion (AA \pm), the cation (AA $^+$) and the anion (AA $^-$). The equilibrium relationship associated to each reaction is:

$$K_D = \frac{[\text{AA}\pm]}{[\text{AA}]} \quad (9)$$

$$K_{A1} = \frac{[\text{AA}^-][\text{H}^+]}{[\text{AA}\pm]} \quad (10)$$

$$K_{A2} = \frac{[\text{AA}\pm][\text{H}^+]}{[\text{AA}^+]} \quad (11)$$

The value of K_D is so high the neutral non-ionized form is virtually inexistent, so it will no longer be considered. The fraction of ionized species of each reaction, excluding the first one, is given by the Henderson–Hasselbalch equations:

$$\alpha_{\text{AA}^-}(\text{pH}) = \frac{10^{\text{pH}-\text{p}K_{A1}}}{1 + 10^{\text{pH}-\text{p}K_{A1}}} \quad (12)$$

$$\alpha_{\text{AA}^+}(\text{pH}) = \frac{1}{1 + 10^{\text{pH}-\text{p}K_{A2}}} \quad (13)$$

Thus, the fraction of each species of amino acid is simply:

$$\phi_{\text{AA}}(\text{pH}) = [1 - \alpha_{\text{AA}^-}(\text{pH})][1 - \alpha_{\text{AA}^+}(\text{pH})] \quad (14)$$

$$\phi_{\text{AA}\pm}(\text{pH}) = \alpha_{\text{AA}^-}(\text{pH})\alpha_{\text{AA}^+}(\text{pH}) \quad (15)$$

$$\phi_{\text{AA}^-}(\text{pH}) = \alpha_{\text{AA}^-}(\text{pH})[1 - \alpha_{\text{AA}^+}(\text{pH})] \quad (16)$$

$$\phi_{\text{AA}^+}(\text{pH}) = [1 - \alpha_{\text{AA}^-}(\text{pH})]\alpha_{\text{AA}^+}(\text{pH}) \quad (17)$$

Eqs. (14) or (15) can be used in Eqs. (4) and (5). The choice between them is completely arbitrary [24], and so the zwitterionic species is chosen for numerical reasons, because absolute values of the fraction calculated through Eq. (14) is vanishing small, several orders of magnitude smaller than absolute values of the fraction evaluated through Eq. (15).

2.2. Extended Pitzer model

To account for liquid phase non-ideality, the extended Pitzer model developed by Pessôa Filho and Maurer [26] was applied. This formulation is more adequate to application to amino acid solutions, wherein the electrically neutral form is a zwitterionic one. For a solute i in an electrolyte aqueous solution, the natural logarithm of the activity coefficient is given by the sum of a term due to long range interactions and one due to short range interactions:

$$\ln \gamma_i = \ln \gamma_i^{\text{LR}} + \ln \gamma_i^{\text{SR}} \quad (18)$$

The long range term is given by:

$$\ln \gamma_i^{\text{LR}} = -A_\phi z_i^2 \left[\frac{2}{b} \ln(1 + b\sqrt{I}) + \frac{\sqrt{I}}{1 + b\sqrt{I}} \right] \quad (19)$$

wherein:

$$A_\phi = \frac{1}{3} \left(\frac{e^2}{\epsilon_0 \epsilon_r k_B T} \right)^{3/2} \frac{(2d_s N_A)^{1/2}}{8\pi} \quad (20)$$

wherein ϵ_0 is the vacuum permittivity ($\epsilon_0 = 8.85419 \times 10^{-12} \text{ A}^2 \text{ s}^4 \text{ kg}^{-1} \text{ m}^{-3}$), ϵ_r is the dielectric constant, d_s is the solvent density, N_A is Avogadro's number ($N_A = 6.0221415 \times 10^{23}$), e is the electronic charge ($e = 1.60218 \times 10^{-19} \text{ As}$), k_B is Boltzmann constant ($k_B = 1.3806503 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$), z_i is the charge of species i

and b is a universal parameter equal to $1.2 \text{ kg}^{1/2} \text{ mol}^{-1/2}$ [25]. The short range term is given by:

$$\ln \gamma_i^{\text{SR}} = 2 \sum_{j \neq s} \lambda_{ij}(I) m_j + 3 \sum_{j \neq s} \sum_{k \neq s} \Lambda_{ijk} m_j m_k - z_i^2 M_s \sum_{j \neq s} \sum_{k \neq s} \lambda_{ij}^{(1)} \frac{1}{a^2 I^2} \times \left[1 - \left(1 + a\sqrt{I} + \frac{a^2 I}{2} \right) \exp(-a\sqrt{I}) \right] m_j m_k \quad (21)$$

wherein M_s is the molar mass of the solvent in kg mol^{-1} , a is another universal parameter equal to $2.0 \text{ kg}^{1/2} \text{ mol}^{-1/2}$, Λ_{ijk} is the three-body interaction coefficient, and λ_{ij} is the two body interaction coefficient, which depends upon ionic strength:

$$\lambda_{ij}(I) = \lambda_{ij}^{(0)} + \lambda_{ij}^{(1)} \frac{2}{a^2 I^2} [1 - (1 + a\sqrt{I}) \exp(-a\sqrt{I})] \quad (22)$$

The ionic strength is defined as:

$$I = \frac{1}{2} \sum_i m_i z_i^2 \quad (23)$$

For the zwitterionic form, the long range term is null since zwitterion has no net charge. Two other simplifications can be made: the first one consists in neglecting three-body interactions ($\Lambda_{ijk} = 0$, for any i, j and k) and the second one in neglecting the dependence upon ionic strength in two-body interactions ($\lambda_{ij}^{(1)} = 0$, for any i and j). The dependence of the parameter due to two-body interactions upon temperature has been correlated by the proposed equation by Pitzer and Peiper [27] with the minimum of adjustable parameters (and for which a theoretical justification can be provided):

$$\lambda_{ij}(T) = C_{ij}^{(1)} + \frac{C_{ij}^{(2)}}{T} \quad (24)$$

Thereby, one can rewrite Eq. (4) using the extended Pitzer model to account for liquid phase non-ideality due to the presence of electric charges in the solution as:

$$\ln \frac{S(\text{pH})}{S(\text{pI})} = \ln \frac{\phi_{\text{AA}\pm}(\text{pI})}{\phi_{\text{AA}\pm}(\text{pH})} + 2\lambda_{\text{AA}^-} [\phi_{\text{AA}^-}(\text{pI})S(\text{pI}) - \phi_{\text{AA}^-}(\text{pH})S(\text{pH})] + 2\lambda_{\text{AA}^+} [\phi_{\text{AA}^+}(\text{pI})S(\text{pI}) - \phi_{\text{AA}^+}(\text{pH})S(\text{pH})] + 2\lambda_{\text{AA}} [\phi_{\text{AA}}(\text{pI})S(\text{pI}) - \phi_{\text{AA}}(\text{pH})S(\text{pH})] + 2\lambda_{\text{AA}\pm} [\phi_{\text{AA}\pm}(\text{pI})S(\text{pI}) - \phi_{\text{AA}\pm}(\text{pH})S(\text{pH})] \quad (25)$$

The empirical observation of amino acid solubility curves as a function of pH reveals certain symmetry around the isoelectric point. This allows inferring that interactions among the electrically neutral molecules of amino acid and electrically charged molecules are indistinguishable, i.e., an electrically neutral molecule interacts in the same way with a positive or negative charged molecule. This hypothesis implies the following identity *a priori*:

$$\lambda_{\text{AA}^-} = \lambda_{\text{AA}^+} = \lambda_{\text{AA}} = \lambda_{\text{AA}\pm} = \lambda \quad (26)$$

Evidently, if such symmetry is not observed, this hypothesis will not hold. Considering the simplification introduced by Eq. (26), Eq. (25) can be rewritten as:

$$\ln \frac{S(\text{pH})}{S(\text{pI})} = \ln \frac{\phi_{\text{AA}\pm}(\text{pI})}{\phi_{\text{AA}\pm}(\text{pH})} + 2\lambda [S(\text{pI}) - S(\text{pH})] \quad (27)$$

By replacing Eqs. (12)–(17) in Eq. (27), one gets an explicit relation between amino acids solubility and pH:

$$\log \frac{S(\text{pH})}{S(\text{pI})} = \text{pI} - \text{pH} + \log \left[\frac{1 + 10^{\text{pH} - \text{p}K_{\text{A}1}}}{1 + 10^{\text{pI} - \text{p}K_{\text{A}1}}} \right] + \log \left[\frac{1 + 10^{\text{pH} - \text{p}K_{\text{A}2}}}{1 + 10^{\text{pI} - \text{p}K_{\text{A}2}}} \right] + \frac{2}{\ln 10} \lambda [S(\text{pI}) - S(\text{pH})] \quad (28)$$

2.3. A statistical thermodynamic interpretation of interaction parameters

Considering McMillan–Mayer's theory, from the statistical thermodynamic point of view, it is possible to emulate the osmotic pressure generated by a solute in a continuum liquid solvent using the same expressions for non-ideality of gases [28]. Therefore, one can write the natural logarithm of the activity coefficient of the zwitterionic molecule as:

$$\ln \gamma_{\text{AA}\pm} = 2 \sum_j \rho_j B_{j,\text{AA}\pm} \quad (29)$$

wherein $B_{j,\text{AA}\pm}$ is the osmotic second virial coefficient and ρ_j is the number of molecules of component j per volume unity.

Considering that the osmotic second virial coefficient can be related to the potential of mean force of interaction among solutes into a given solvent in the limit of low solute concentrations, one gets [29]:

$$B_{ij} = -\frac{1}{16\pi^2} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} \int_0^{2\pi} \int_0^\pi \int_0^{+\infty} \left[\exp\left(-\frac{W_{ij}(r)}{k_B T}\right) - 1 \right] r^2 \sin \varphi \sin \beta r \, dr \, d\varphi \, d\theta \, d\alpha \, d\beta \, d\gamma \quad (30)$$

Considering the mean force potential as given by the sum of three contributions: hard sphere (W^{HS}), attractive potential due to van der Waals forces (W^{vdW}) and electrostatic interactions (W^{elet}) as suggested by the DLVO theory:

$$W_{ij}(r) = \begin{cases} +\infty, & \text{if } r \leq \sigma \\ z_i z_j e^2 \exp[-\kappa(r - \sigma)] / \epsilon r(1 + \kappa\sigma) - \epsilon \left(\frac{\sigma}{r}\right)^6, & \text{if } r > \sigma \end{cases} \quad (31)$$

wherein z_i is the electric charge of molecule i , z_j is the electric charge of molecule j , e is the electronic charge, σ is the mean van der Waals diameter of molecules i and j , κ is the interaction parameter between molecules i and j .

One can assume that any configuration of an amino acid molecule has the same van der Waals diameter, σ – a reasonable hypothesis since the only difference between them is the charge. We also assumed that the magnitude of van der Waals interaction parameter between zwitterionic form and the others is the same, which also seems reasonable for such similar chemical structures as the ionized forms of an amino acid.

Being molecule i the zwitterion, electrostatic contribution is null since z_i equals zero. Thus, the osmotic second virial coefficient for a system with isotropic interactions can be written as:

$$B_{ij} = \frac{2\pi\sigma^3}{3} \left(1 - \frac{\epsilon}{k_B T} \right) \quad (32)$$

The natural logarithm of zwitterionic activity coefficient becomes:

$$\ln \gamma_{\text{AA}\pm} = \frac{4\pi\sigma^3}{3} \left(1 - \frac{\epsilon}{k_B T} \right) \sum_j \rho_j \quad (33)$$

Table 1
Values of the interaction parameter (λ) and RMSD – Eq. (37) – for amino acids.

Amino acids	λ (kg mol ⁻¹)	RMSD (%)	Experimental data
D,L-Alanine	0.0267	2.0	[8]
D,L-Methionine	0.0935	6.8	[13]
L-Isoleucine	0.234	5.2	[8]
L-Leucine	0.273	9.4	[8]
D,L-Phenylalanine	3.95	2.9	[8]
L-Serine	0.0165	4.6	[8]

Replacing Eq. (33) in Eq. (4), one has:

$$\ln \frac{S(\text{pH})}{S(\text{pI})} = \ln \frac{\phi_{AA\pm}(\text{pI})}{\phi_{AA\pm}(\text{pH})} + \frac{4\pi\sigma^3 N_A \rho}{3} \left(1 - \frac{\varepsilon}{k_B T}\right) [S(\text{pI}) - S(\text{pH})] \quad (34)$$

From the comparison between Eqs. (27) and (34), one has that the physical meaning of parameters established in Eq. (24) should be given by:

$$C_{ij}^{(1)} = \frac{2\pi\sigma^3 N_A \rho}{3} \quad (35)$$

$$C_{ij}^{(2)} = -C_{ij}^{(1)} \frac{\varepsilon}{k_B} \quad (36)$$

3. Results and discussions

3.1. Solubility of amino acids as a function of pH

The roots of Eq. (28) were obtained through the Newton–Raphson method. Parameters $C_{ij}^{(1)}$ and $C_{ij}^{(2)}$ were obtained using the simplex method of Nelder and Mead [30]. The relative mean square deviation (RMSD) was used as an objective function to adjust interaction parameters along with Eq. (28):

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{S_i^{\text{exp}} - S_i^{\text{calc}}}{S_i^{\text{exp}}} \right)^2} \quad (37)$$

wherein S_i^{exp} is the i th experimental solubility and S_i^{calc} is the related calculated solubility by Eq. (28).

Table 1 presents the results for interaction parameter for systems containing amino acids. The values of $\text{p}K_A$ for such systems are presented in Table 2. One can observe that the maximum value of RMSD for such adjustments is less than 10.0% and that the interaction parameter magnitude increases with the amino acid molecular weight. This finding shows that non-ideality is mainly given by an entropic (combinatorial) effect, since it is dependent on the size of the molecules. The results are presented in Figs. 1–6. From them, it is possible to state that the application of the model is a proficuous approach.

Eq. (37) can be used as a criterion for a quantitative comparison among different models. We compared the results of the model herein presented with the ideal solution model and with the NRTL model applied by Tseng et al. [8]. There is an important physical difference between the application of this model and that of NRTL as applied by Tseng et al. [8]. While this model is in McMillan–Mayer ensemble, as well as Pitzer’s model, considering that non-ideality is established by interaction among different amino acids molecules, the NRTL model considers that non-ideality is due to the interaction between the amino acid and the solvent molecule. Table 3 shows the comparison among these models.

While the model herein presented requires only one adjustable parameter, the NRTL equation requires at least two. Even so, as one can see in Table 3, this model correlates experimental data with relatively more accuracy for all the systems studied. For systems

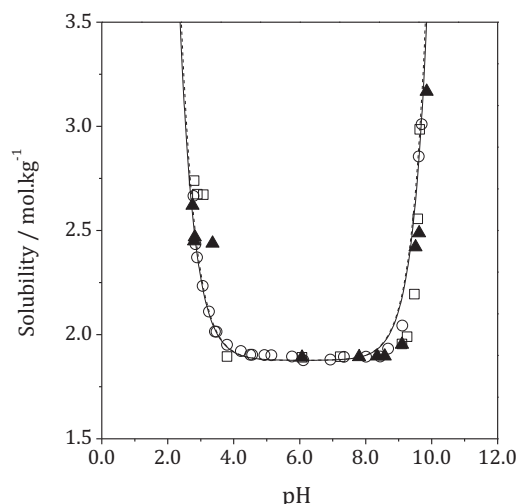


Fig. 1. Solubility of DL-alanine as a function of pH at 298.15 K in aqueous solution. Experimental data (○) [8]; NaCl or HCl solution (▲) [6]; KOH or HNO₃ solution (□) [6]. Modeling: ideal solution (dotted line), this work (continuous line).

such as solutions containing DL-alanine and L-serine, despite the vast literature concerning the non-ideality modeling of those systems, the simplest hypothesis of ideal solution is reasonable when the RMSD is compared among different models.

Table 2
Values of $\text{p}K_A$ for amino acids, peptides and β -lactam compounds.

Compound	T (K)	$\text{p}K_{A1}$	$\text{p}K_{A2}$	Reference
D,L-Alanine	298.15	2.35	9.87	[8]
D,L-Methionine	298.15	2.12	9.28	[13]
L-Isoleucine	298.15	2.32	9.76	[8]
L-Leucine	298.15	2.33	9.74	[8]
D,L-Phenylalanine	298.15	2.58	9.24	[8]
L-Serine	298.15	2.21	9.15	[8]
Tyrosine ^a	298.15	2.20	9.11	[32]
Glycine	298.15	2.34	9.60	[32]
Dyglycine	298.15	3.12	8.17	[32]
Triglycine	298.15	3.26	7.91	[32]
Tetraglycine	298.15	3.05	7.75	[32]
Pentaglycine	298.15	3.05	7.70	[32]
Hexaglycine	298.15	3.05	7.60	[32]
Ampicillin	283.1	2.27	7.66	[22]
	288.1	2.24	7.61	[22]
	293.1	2.14	7.45	[22]
	298.1	2.14	7.31	[22]
6-Aminopenicillanic acid	283.1	1.98	5.34	[22]
	288.1	2.04	5.25	[22]
	293.1	2.1	5.18	[22]
	298.1	2.20	4.83	[22]

^a The $\text{p}K_{A3}$ value of tyrosine is 10.07 according to Greenstein and Winitz [32].

Table 3
Values of RMSD – Eq. (37) – for amino acids applying the models: ideal solution, this work and NRTL. Experimental solubility data obtained by Tseng et al. [8].

Amino acid	Model		
	Ideal solution (%)	This work (%)	NRTL ^a (%)
D,L-Alanine	2.4	2.0	2.2
L-Isoleucine	24.9	5.2	6.3
L-Leucine	32.3	9.4	13.3
D,L-Phenylalanine	62.5	2.9	3.0
L-Serine	6.0	4.6	5.1

^a NRTL model applied with parameters estimated by Tseng et al. [8].

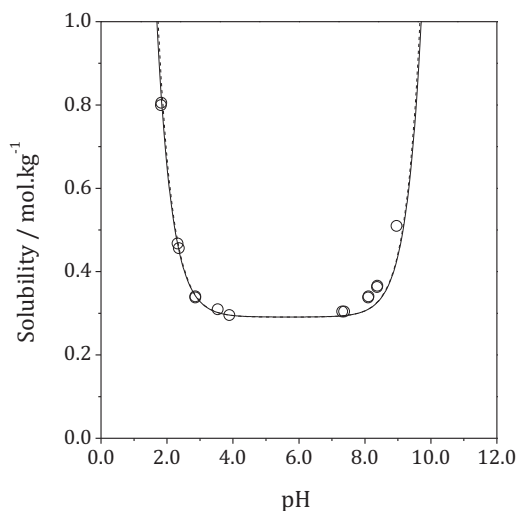


Fig. 2. Solubility of DL-methionine as a function of pH at 303.0 K in aqueous solution. Experimental data (○) [13]. Modeling: ideal solution (dotted line), this work (continuous line).

The model must be slightly modified to account for the presence of amino acids with ionizable residues. With one more ionizable group, there will be eight different configurations that can be assumed by an amino acid molecule. Therefore, Eq. (25) would have eight, instead of four, interaction parameters: one parameter for the interaction of each amino acid configuration with the zwitterionic form. If we adopt a similar hypothesis as that adopted in Eq. (26), Eq. (28) can thus be rewritten as:

$$\log \frac{S(\text{pH})}{S(\text{pI})} = \text{pI} - \text{pH} + \log \left[\frac{1 + 10^{\text{pH} - \text{p}K_{A1}}}{1 + 10^{\text{pI} - \text{p}K_{A1}}} \right] + \log \left[\frac{1 + 10^{\text{pH} - \text{p}K_{A2}}}{1 + 10^{\text{pI} - \text{p}K_{A2}}} \right] + \log \left[\frac{1 + 10^{\text{pH} - \text{p}K_{A3}}}{1 + 10^{\text{pI} - \text{p}K_{A3}}} \right] + \frac{2}{\ln 10} \lambda [S(\text{pI}) - S(\text{pH})] \quad (38)$$

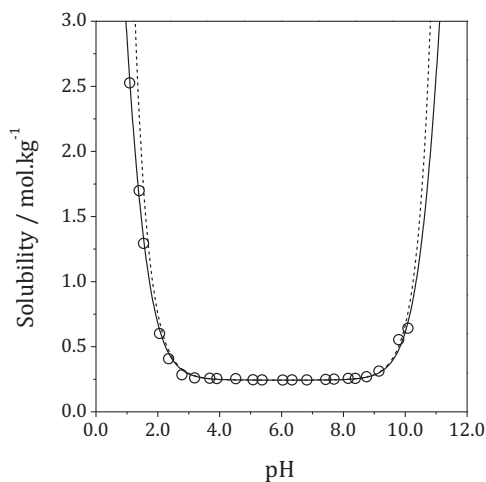


Fig. 3. Solubility of L-isoleucine as a function of pH at 298.15 K in aqueous solution. Experimental data (○) [8]. Modeling: ideal solution (dotted line), this work (continuous line).

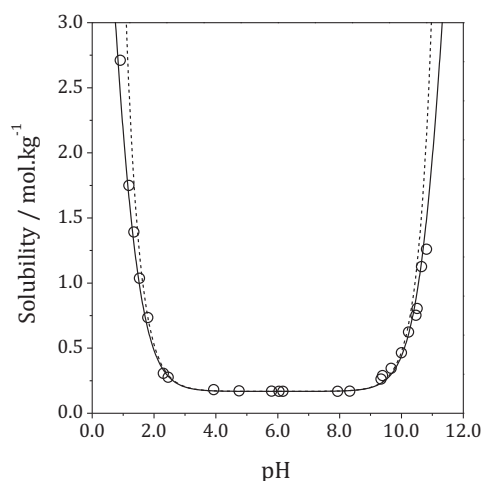


Fig. 4. Solubility of L-leucine as a function of pH at 298.15 K in aqueous solution. Experimental data (○) [8]. Modeling: ideal solution (dotted line), this work (continuous line).

There are few experimental data available in the open literature of systems containing amino acids with three ionizable groups in chain. Here, we used the tyrosine solubility data published by Hitchcock [15]. To the best of our knowledge, in that article, Hitchcock presented the very first modeling for an amino acid solubility curve as a function of pH, considering ideal solution. Although tyrosine has an ionizable R-group, it commonly appears in the biochemical specialized literature as nonpolar amino acid since its $\text{p}K_{A3}$ is too high. In Fig. 7, we show how important is to consider the tyrosine R-group as ionizable even when considering only ideal solution ($\lambda_{ij} = 0$). And it is interesting to note that considering the tyrosine R-group as an ionizable one does improve solid-liquid equilibrium description.

3.2. Solubility of glycine and its oligopeptides as a function of pH

The solubility of small peptides with only terminal groups, such as ionizable ones was also investigated. Table 4 presents the values

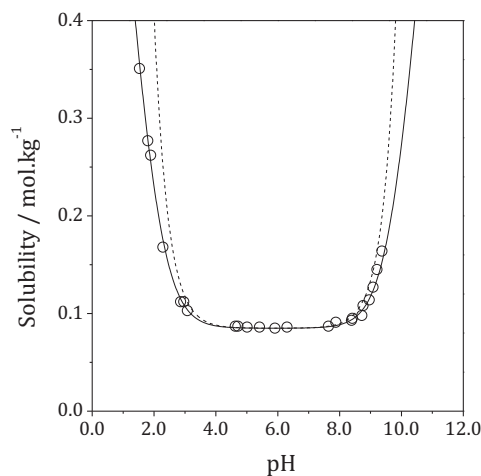


Fig. 5. Solubility of DL-phenylalanine as a function of pH at 298.15 K in aqueous solution. Experimental data (○) [8]. Modeling: ideal solution (dotted line), this work (continuous line).

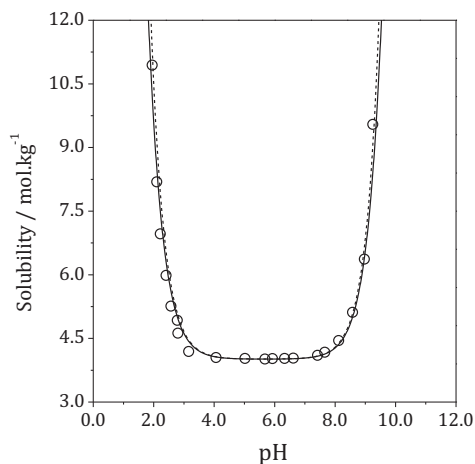


Fig. 6. Solubility of L-serine as a function of pH at 298.15 K in aqueous solution. Experimental data (○) [8]. Modeling: ideal solution (dotted line), this work (continuous line).

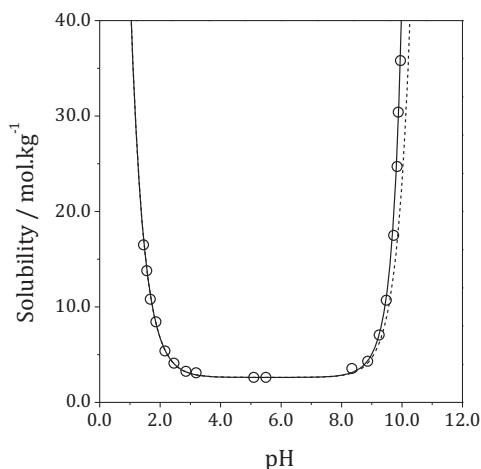


Fig. 7. Solubility of tyrosine as a function of pH at 298.15 K in aqueous solution. Experimental data (○) [15]. Modeling: considering R-group as non-ionizable (dotted line), considering R-group as ionizable (continuous line).

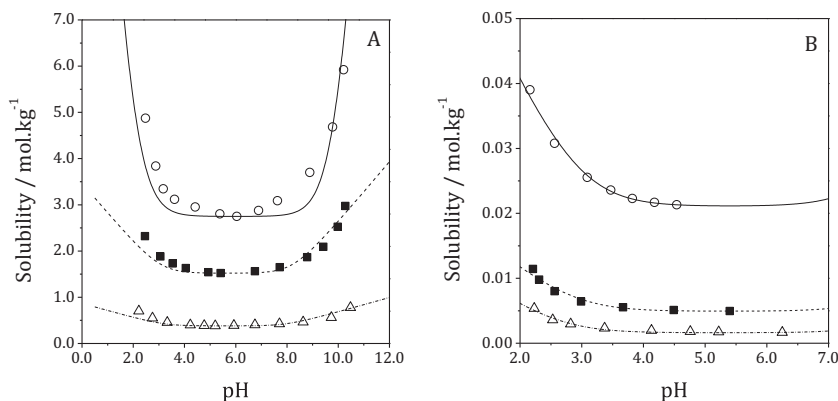


Fig. 8. Solubility of glycine and its oligopeptides as a function of pH. Experimental data [31]: (○) glycine (A), tetraglycine (B); (■) dyglycine (A), pentaglycine (B); and (△) triglycine (A), hexaglycine (B). Our model: (continuous line) glycine (A), tetraglycine (B); (dashed line) dyglycine (A), pentaglycine (B); and (dashed-dotted line) triglycine (A), hexaglycine (B).

Table 4

Values of interaction parameter (λ) and RMSD – Eq. (37) – for glycine and its oligopeptides as a function of pH. Experimental solubility data obtained by Lu et al. [31].

Compound	λ (kg mol ⁻¹)	RMSD (%)
Glycine	0.200	11.0
Dyglycine	3.26	6.9
Triglycine	13.8	9.8
Tetraglycine	93.6	1.7
Pentaglycine	238.8	4.7
Hexaglycine	258.2	7.9

of interaction parameter λ_{ij} and of RMSD – defined in Eq. (37) – for aqueous solutions containing glycine or its oligopeptides (up to hexaglycine). The values of pK_A are also presented in Table 2. Fig. 8 allows observing that the model can accurately describe solubility curves. It is interesting that, as similar as for simple amino acids, the magnitude of interaction parameter λ_{ij} increases with the increment in the peptide chain size. Considering the interpretation of the parameter presented earlier, one can write:

$$\lambda = \frac{2\pi\sigma^3}{3} N_A \rho \left(1 - \frac{\varepsilon}{k_B T} \right) \quad (39)$$

Assuming that ε should increase as the peptide nonpolar chain increases, we would expect λ to decrease with the increment in peptide chain size, since there would be an increment in hydrophobic interaction represented by the attractive parameter ε . However, the value of λ increases, which means that the effect of van der Waals diameter σ is predominant.

3.3. Solubility of β -lactam compounds as a function of pH and temperature

Similarly to amino acids, β -lactam compounds present one amine and one carboxylic acid that participate in protonation equilibria. This common feature turns the developed modeling for amino acids applicable to β -lactam compounds. Fig. 9 shows the chemical structures of two β -lactam compounds: the antibiotic ampicillin and its precursor 6-aminopenicillanic acid, for which solubility data as a function of pH and temperature are available [22]. In this case, one can apply Eq. (28) by simultaneously adjusting parameters $C_{ij}^{(1)}$ and $C_{ij}^{(2)}$, which are related to σ and ε according to Eqs. (35) and (36). The regression was performed using Nelder and

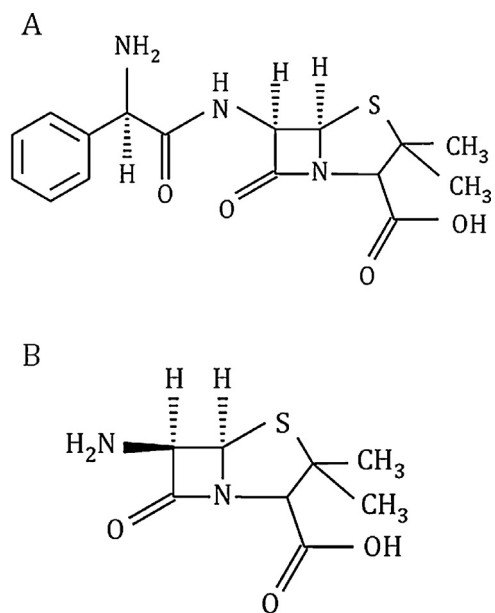


Fig. 9. Chemical structure of ampicillin (A) and 6-aminopenicillanic acid (B).

Mead [30] simplex method with the objective function of global relative mean square deviation (GRMSD):

$$\text{GRMSD} = \sqrt{\frac{1}{M \times N} \sum_{j=1}^M \sum_{i=1}^N \left(\frac{S_{i,j}^{\text{exp}} - S_{i,j}^{\text{calc}}}{S_{i,j}^{\text{exp}}} \right)^2} \quad (40)$$

wherein N is the number of experimental data point at a given temperature and M is the number of sets of experimental data at different temperatures. The values of pK_A used for ampicillin and 6-aminopenicillanic acid are listed in Table 2.

Table 5 presents values of parameter σ and ε adjusted as well as the value of GRMSD defined in Eq. (40). The adjusted values are coherent, as the values of σ are internally consistent, considering the size of ampicillin and 6-aminopenicillanic acid molecules. The values of ε/k_B show an analogous coherence.

Accounting for non-ideality for β -lactam solutions is necessary to describe the solubility experimental data, as can be viewed in Figs. 10 and 11. In both cases, the dependence of solubility on temperature is weak, but non-negligible. The values of ε/k_B are related to the deviation (positive or negative) from ideality: if ε/k_B is larger than the temperature of the experimental data set, the activity coefficient of zwitterionic species is lower than the unity, resulting in a negative deviation from ideality.

Table 6 presents the comparison among three models: ideal solution, Khoshkbarchi and Vera [11] model and the model herein presented. The value of GRMSD for the model herein presented can be seen to be lower than that obtained by applying the ideal solution model, and even lower than the Khoshkbarchi and Vera model [11] – which requires five adjustable parameters.

Table 5

Values of parameters σ and ε , and values of GRMSD – Eq. (40) – for systems containing ampicillin and 6-aminopenicillanic acid. Experimental solubility data obtained by Santana et al. [22].

Compound	σ (Å)	ε/k_B (K)	GRMSD (%)
Ampicillin	5.25	299.8	8.2
6-Aminopenicillanic acid	2.68	275.6	12.0

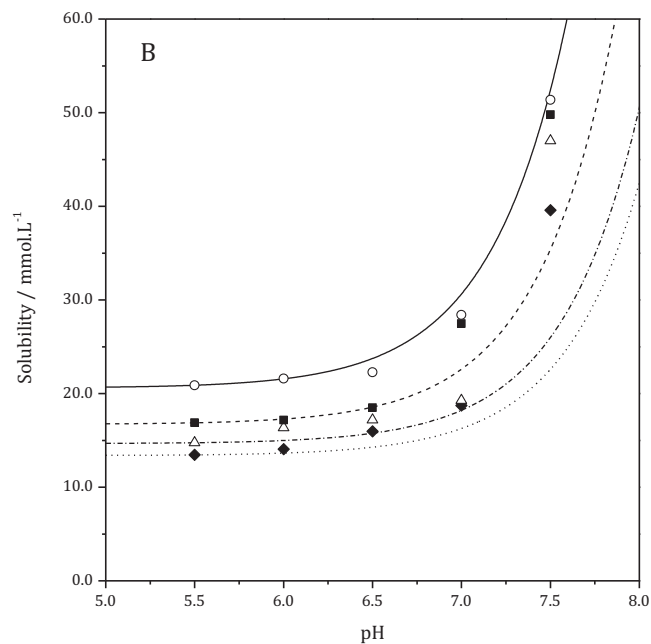
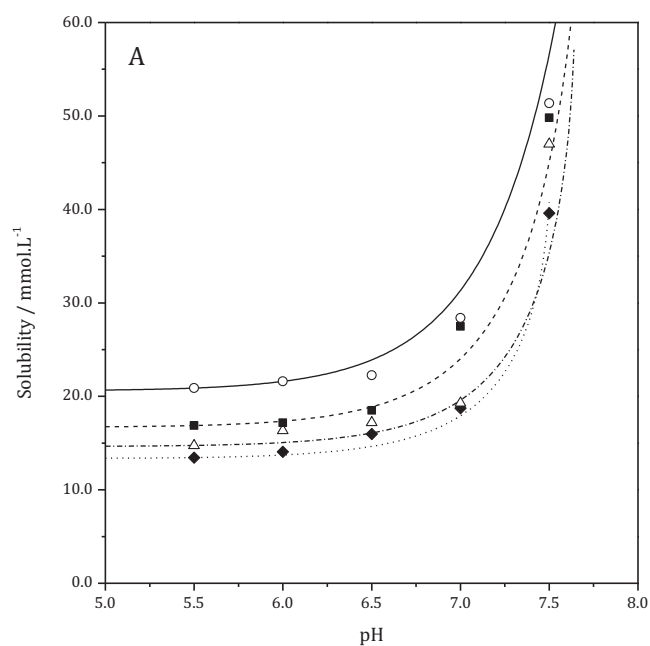


Fig. 10. Solubility of ampicillin as a function of pH in aqueous solution. Experimental data [22]: (◆) 283.06 K, (△) 288.01 K, (■) 292.95 K and (○) 298.03 K. Modeling: dotted line – 283.06 K, dot-dashed line – 288.01 K, dashed line – 292.95 K, continuous line – 298.03 K. (A) This work; (B) ideal solution.

Table 6

Values of GRMSD – Eq. (40) – for different models applied to systems containing ampicillin and 6-aminopenicillanic acid. Experimental solubility data obtained by Santana et al. [22].

Model	Compound	
	Ampicillin (%)	6-Aminopenicillanic acid (%)
Ideal solution	16.6	148.4
Khoshkbarchi and Vera [11] ^a	8.6	55.6
This work	8.2	12.0

^a Model by Khoshkbarchi and Vera [11] applied with parameters estimated by Santana et al. [22].

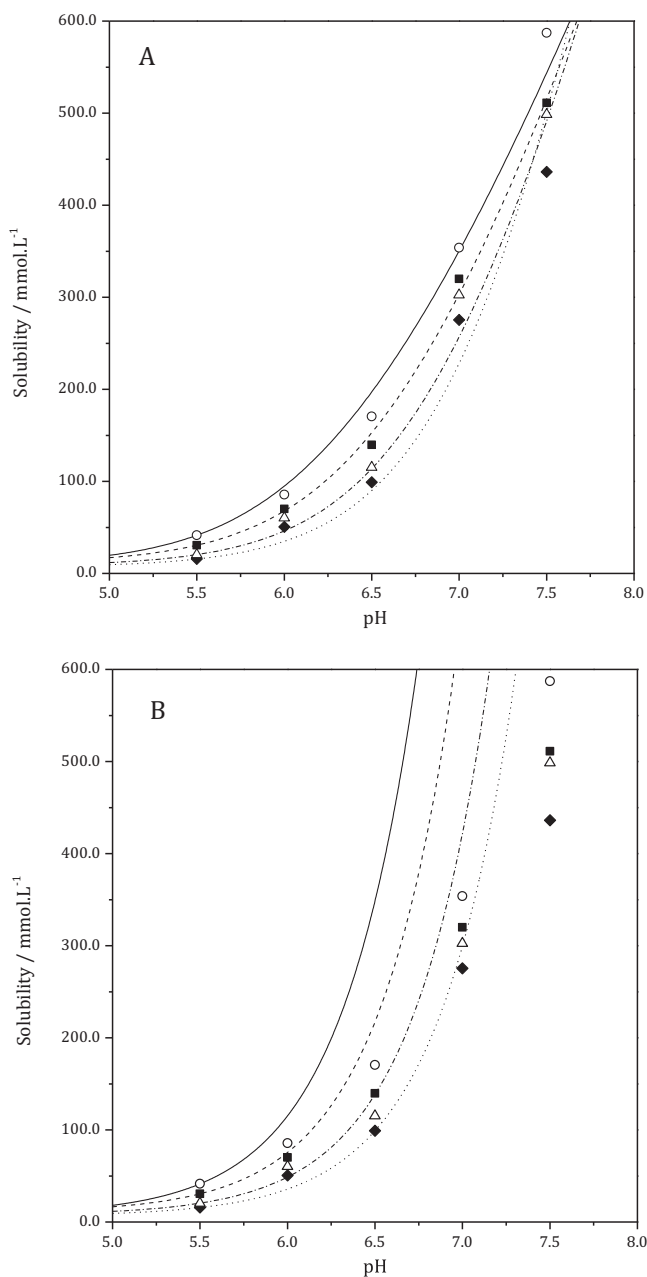


Fig. 11. Solubility of 6-aminopenicillanic acid as a function of pH in aqueous solution. Experimental data [22]: (◆) 283.06 K, (△) 288.01 K, (■) 292.95 K and (○) 298.03 K [22]. Modeling: dotted line – 283.06 K, dot-dashed line – 288.01 K, dashed line – 292.95 K, continuous line, 298.03 K. (A) This work; (B) ideal solution.

4. Conclusions

A new model for describing the solubility curve of amino acids and β -lactam compounds as a function of pH is presented. This

model requires less adjustable parameters and allows a better correlation of the experimental data than traditional approaches. The applicability of this model in the correlation of several sets of solubility experimental data is stressed. An important observation concerns the combinatorial entropic effect, which might explain the increasing non-ideality with increment of molecular size. This fact allied to the literature review shows that the use of very complicated models to describe non-ideality in systems such as the ones studied here may not be the best way to develop thermodynamic modeling for engineering purposes.

Acknowledgments

The authors gratefully acknowledge the financial support from the Brazilian agencies FAPESP (processes 2008/11232-1 and 2011/22070-5), CNPq and CAPES.

References

- [1] J.G. Kirkwood, *J. Chem. Phys.* 2 (1934) 351–361.
- [2] K.K. Nass, *AIChE J.* 34 (1988) 1257–1266.
- [3] C.C. Chen, Y. Zhu, L.B. Evans, *Biotechnol. Prog.* 5 (1989) 111–118.
- [4] R. Rodríguez-Raposo, L. Fernández-Mérida, M.A. Esteso, *J. Chem. Thermodyn.* 26 (1994) 1121–1128.
- [5] M.K. Khoshkbarchi, J.H. Vera, *AIChE J.* 42 (1996) 2354–2364.
- [6] A.A. Pradhan, J.H. Vera, *Fluid Phase Equilib.* 152 (1998) 121–132.
- [7] X. Xu, S.P. Pinho, E.A. Macedo, *Ind. Eng. Chem. Res.* 43 (2004) 3200–3204.
- [8] H.C. Tseng, C.Y. Lee, W.L. Weng, I.M. Shiah, *Fluid Phase Equilib.* 285 (2009) 90–95.
- [9] R.B. Gupta, R.A. Heidemann, *AIChE J.* 36 (1990) 333–341.
- [10] S.P. Pinho, C.M. Silva, E.A. Macedo, *Ind. Eng. Chem. Res.* 33 (1994) 1341–1347.
- [11] M.K. Khoshkbarchi, J.H. Vera, *Ind. Eng. Chem. Res.* 35 (1996) 4319–4327.
- [12] B.H. Park, K.P. Yoo, C.S. Lee, *Fluid Phase Equilib.* 212 (2003) 175–182.
- [13] D. Fuchs, J. Fischer, F. Tumakaka, G. Sadowski, *Ind. Eng. Chem. Res.* 45 (2006) 6578–6584.
- [14] S. Seyfi, G. Pazuki, S.F. Aghamiri, M. Behesti, *Fluid Phase Equilib.* 287 (2009) 15–22.
- [15] D. Hitchcock, *J. Gen. Physiol.* 6 (1924) 747–757.
- [16] J.G. Kirkwood, *Chem. Rev.* 24 (1939) 233–251.
- [17] G.M. Wilson, *J. Am. Chem. Soc.* 86 (1964) 127–130.
- [18] H. Renon, J.M. Prausnitz, *AIChE J.* 14 (1968) 135–144.
- [19] B.L. Larsen, P. Rasmussen, A. Fredenslund, *Ind. Eng. Chem. Res.* 26 (1987) 2274–2286.
- [20] L.A. Bromley, *AIChE J.* 19 (1973) 313–320.
- [21] M.K. Khoshkbarchi, J.H. Vera, *AIChE J.* 42 (1996) 249–258.
- [22] M. Santana, M.P.A. Ribeiro, G.A. Leite, R.L.C. Giordano, R.C. Giordano, S. Mattedi, *AIChE J.* 56 (2010) 1578–1583.
- [23] J.B.G. Dalrup, C. Held, G. Sadowski, G. Schembecker, *Ind. Eng. Chem. Res.* 50 (2011) 3503–3509.
- [24] L.F.M. Franco, P.A. Pessôa Filho, *Fluid Phase Equilib.* 306 (2011) 242–250.
- [25] J.M. Prausnitz, R.N. Lichtenthaler, E.G. Azevedo, *Molecular Thermodynamics of Fluid-Phase Equilibria*, 3rd ed., Prentice Hall, 1999.
- [26] P.A. Pessôa Filho, G. Maurer, *Fluid Phase Equilib.* 269 (2008) 25–35.
- [27] K.S. Pitzer, J.C. Peiper, *J. Phys. Chem. Ref. Data* 13 (1984) 1–101.
- [28] T.L. Hill, *An Introduction to Statistical Thermodynamics*, Dover, 1986.
- [29] B.L. Neal, D. Asthagiri, O.D. Velev, A.M. Lenhoff, E.W. Kaler, *J. Cryst. Growth* 196 (1999) 377–387.
- [30] J.A. Nelder, R.A. Mead, *Comp. J.* 7 (1965) 308–313.
- [31] J. Lu, X.J. Wang, X. Yang, C.B. Ching, *J. Chem. Eng. Data* 51 (2006) 1593–1596.
- [32] J.P. Greenstein, M.M. Winitz, *Chemistry of the Amino Acids*, John Wiley & Sons, New York, 1961.