Solubility Models for Amino Acids and Antibiotics

An approach is presented for modeling the effects of temperature and pH on the solubility of amino acids in water. The amino acids studied are alanine, amino-butyric acid, glycine, hydroxy-proline, proline, serine, threonine, and valine. The data employed are activity coefficients and solubilities of the amino acids in neutral water and the dissociation constants for the various amino-acid ionization reactions. Activity coefficients are correlated with the modified UNIFAC group contribution model and with new glycine and proline groups being introduced.

A similar approach is presented for modeling the solubilities of certain antibiotics in mixed nonaqueous solvents. Fusion temperature data and solubilities of the antibiotics in pure organic solvents are used to deduce modified UNIFAC energy interaction parameters between new large antibiotic groups and the standard alkane (CH_2), alcohol (OH), aromat (ACH), and ester (CCOO) groups.

Ram B. Gupta
Robert A. Heidemann
Department of Chemical and Petroleum Engineering
University of Calgary
Calgary, Alberta, T2N 1N4, Canada

Introduction

In a typical case, the cost involved in bioseparations may be as high as ninety percent of the total cost of manufacturing. Unfortunately, this area has not been given much research attention and bioseparation units are most often designed empirically rather than on the basis of rational information.

The prediction of activity coefficients for amino acids has had only recent attention. Orella and Kirwan (1987), in their unpublished work, have made an attempt to predict solubilities of several amino acids (glycine, alanine, phenylalanine, valine, leucine, and iso-leucine). They considered the activity coefficient to be made up of terms due to chemical interactions (from UNIFAC) and due to electrostatic interactions (from the extended Kirkwood theory).

Nass (1988) has correlated amino acid activity coefficient and solubility data. She assumed that the activity coefficient is a product of terms due to chemical reaction equilibrium and due to physical interaction. Wilson's equation was used for the physical activity coefficient, with Bondi's volumes (Bondi, 1968) inserted for the pure-component liquid volume ratios. Activity coefficients for alanine, serine, and threonine in water, and solubilities of phenylalanine, tyrosine, and diiodotyrosine in water have been correlated. The correlations are in good agreement with the experimental data. The number of parameters regressed varied from three to ten.

Correspondence concerning this paper should be addressed to R. A. Heidemann.

In this work, an attempt has been made to build a predictive model for the activity coefficients of amino acids. Amino acids in the study are alanine, amino-butyric acid, glycine, hydroxy-proline, proline, serine, threonine, and valine. These are eight of the twenty most common amino acids. The effect of pH and temperature on amino acid solubility are considered.

We have also explored the possibility of constructing a predictive model for the solubility in mixed solvents of six antibiotics—anisomycin, carbomycin A, chloramphenicol, chloramphenicol palmitate, griseofulvin, and hygromycin A. The selection of these antibiotics is based on their commercial importance and the availability of solubility, melting temperature, and chemical structure data.

Drugs such as these antibiotics are commonly extracted from a fermentation broth with a solvent that is immiscible with water. Recovery of the product is achieved by dilution with additional solvents in which the solid has a lower solubility.

Model Development

Amino acids

When the amino acids are dissolved in water, they are almost completely converted to the zwitterion form and/or to one of the ions with a net charge. In treating the amino acids, we have assumed that activity coefficients of all the ionic species are equal and can be evaluated from modified UNIFAC using the mole fraction in the UNIFAC model as if no ionization occurred

at all. However, an accurate representation of solubility variation with pH requires attention to the ionization reactions.

Dissociation of Amino Acids. Amino acids take several ionic forms when dissolved in water. The following reactions occur:

$$NH_2RCOOH = NH_3{}^+RCOO^- \quad K_D \tag{1}$$

$$NH_3{}^+RCOOH = H^+ + NH_3{}^+RCOO^- \quad K_1 \tag{2}$$

$$NH_3{}^+RCOO^- = H^+ + NH_2RCOO^- \quad K_2 \tag{3}$$

$$H_2O = H^+ + OH^- \quad K_w \tag{4}$$

Reaction 1 shows formation of the neutral zwitterion with its charge distribution. The zwitterion can accept a proton, as shown in reaction 2 to form the positively charged species on the lefthand side of the equation. It can also donate a proton as shown in reaction 3 to produce a negatively charged species. The dissociation equilibrium constants K_D, K_1, and K_2 are given by

$$K_D = \frac{[NH_3{}^+RCOO^-]}{[NH_2RCOOH]} \tag{5}$$

$$K_1 = \frac{[H^+]\,[NH_3{}^+RCOO^-]}{[NH_3{}^+RCOOH]} \tag{6}$$

$$K_2 = \frac{[H^+][NH_2RCOO^-]}{[NH_3{}^+RCOO^-]} \tag{7}$$

where the brackets indicate molality units.

Greenstein and Winitz (1961) report that the value of K_D for aliphatic amino acids is on the order of 10^5 to 10^6. This means that the amount of the uncharged amino acid in the aqueous solution is negligible. The amino acids studied for the effect of pH, in this work, have a very high value of K_D; hence it is assumed that the amino acid, when dissolved in water, exists mainly as the zwitterion and the two net charged species.

On dissolution of the amino acid in neutral water, the solution becomes acidic, since the amino acid behaves like a weak acid. The pH of the solution can be calculated from a mass balance.

Suppose that n_o moles of the amino acid are added to 1 kg of water. Let reaction extents ϵ_1, ϵ_2, and ϵ_3 be the moles of H^+ produced in the reactions 2, 3, 4, respectively. Then, supposing ϵ_1 and ϵ_2 are very small in comparison with n_o, the equilibrium equations are.

$$K_1 = \frac{(\epsilon_1 + \epsilon_2 + \epsilon_3)n_o}{-\epsilon_1} \tag{8}$$

$$K_2 = \frac{(\epsilon_1 + \epsilon_2 + \epsilon_3)(\epsilon_2)}{n_o} \tag{9}$$

$$K_w = (\epsilon_1 + \epsilon_2 + \epsilon_3)\,(\epsilon_3) \tag{10}$$

These three equations can be solved for ϵ_1, ϵ_2, and ϵ_3, yielding for the hydrogen ion molality,

$$[H^+]^2 = (\epsilon_1 + \epsilon_2 + \epsilon_3)^2 = \frac{K_w K_1 + n_o K_1 K_2}{K_1 + n_o} \tag{11}$$

From Eq. 11, for low n_o

$$[H^+] = (K_w)^{1/2} = 10^{-7} \tag{12}$$

that is, $pH = -\log_{10}[H^+] = 7$. On the other hand, as n_o becomes larger

$$[H^+] = (K_1 K_2)^{1/2} \tag{13}$$

and

$$pH = \frac{-(\log_{10} K_1 + \log_{10} K_2)}{2} \tag{14}$$

or

$$pH = \frac{pK_1 + pK_2}{2} \tag{15}$$

The value of pH given by Eq. 15 is called the isoelectric point for the amino acid.

Typical values for pK_1 and pK_2 are 2.3 and 9.7, respectively. The corresponding pH in all but very dilute solutions of amino acids is pH = 6.0. When the molality n_o is only 5×10^{-3} the pH is 6.15.

Amino Acid Solubility. Using x_A for undissociated amino acid actually present in a saturated solution, solid-liquid equilibrium of the amino acid is given by;

$$x_A \gamma_A f^o_{A_L} = f_{A_S} \tag{16}$$

The zwitterion is in chemical equilibrium with undissociated A. Hence, from the definition of the dissociation equilibrium constant and assumed equal activity coefficients of all amino acid ions:

$$K_D = \frac{m_{A_\pm}}{m_A} = \frac{x_{A_\pm}}{x_A} \tag{17}$$

Therefore, at equilibrium,

$$x_{A_\pm}\gamma_A\left[\frac{f^o_{A_L}}{K_D}\right] = f_{A_S} \tag{18}$$

The amino acids decompose without melting and f^o_{AL} always refers to a hypothetical liquid state. It is convenient to think of $(f^o_{A_L}/K_D)$, as being $f^o_{A_\pm L}$, the fugacity of the zwitterion in a hypothetical pure-liquid standard state. Then

$$x_{A_\pm}\gamma_A = \frac{f_{a_s}}{f^o_{A_\pm L}} \tag{19}$$

Because the actual mole fraction of the zwitterion depends on the solution pH, the solubility of the amino acid in water will vary as pH is varied.

If x^o_A is the mole fraction of the amino acid calculated as if there were no ionization, then

$$x^o_A = x_{A_\pm} + x_{A_+} + x_{A_-} \tag{20}$$

(The mole fraction of the undissociated amino acid is assumed negligible.) Using the equilibrium constants K_1 and K_2 of Eqs. 2 and 3 yields

$$x_{A_+} = \frac{[H^+]x_{A_\pm}}{K_1} \tag{21}$$

and

$$x_{A_-} = \frac{K_2 x_{A_\pm}}{[H^+]} \tag{22}$$

Then

$$x_A^o = x_{A_\pm}\left[1 + \frac{[H^+]}{K_1} + \frac{K_2}{[H^+]}\right] \tag{23}$$

The equation for the solid-solution equilibrium is then

$$x_A^o \gamma_A = \left[\frac{f_{A_S}}{f^o_{A_{\pm L}}}\right]\left[1 + \frac{[H^+]}{k_1} + \frac{K_2}{[H^+]}\right] \tag{24}$$

When γ_A is known as a function of amino acid mole fraction (on an un-ionized basis) and when $(f_{A_S}/f^o_{A_{\pm L}})$ is known as a function of temperature, Eq. 24 can be used to determine the solubility of the amino acid in solutions of varying pH.

In this study, γ_A data from osmotic coefficient measurements have been used to determine the modified UNIFAC energy parameters. Solubility information for the amino acids in neutral water at various temperatures has been used to correlate $(f_{A_S}/f^o_{A_{\pm L}})$ in the form

$$\ln\left[\frac{f_{A_S}}{f^o_{A_{\pm L}}}\right] = A' - \frac{B'}{T} \tag{25}$$

by using Eq. 24 with $[H^+]$ calculated from Eq. 11.

Antibiotics

Equilibrium between a pure solid antibiotic and the species in solution requires equal fugacities; that is,

$$f_L = f_S \tag{26}$$

The fugacity of a substance in solution can be calculated from an activity coefficient model and a standard state fugacity, so that at saturation (when $x = x^s$ and $\gamma = \gamma^s$),

$$\gamma^s x^s f^o_L = f_S \tag{27}$$

and therefore

$$\gamma^s x^s = f_S / f^o_L \tag{28}$$

Assuming that the righthand side of Eq. 3 is known, the equation can be solved for γ^S.

The solubilities of the antibiotics in most solvents is very low and experiments that might lead to information on the variation of the antibiotic activity coefficient with mole fraction in a given solvent are generally infeasible. However, it is possible to obtain the activity coefficient at saturation, γ^s, and the respective saturation mole fraction of a given antibiotic in a number of chemically similar solvents. Information of this nature is usable to determine UNIFAC energy interaction parameters between antibiotic groups and standard UNIFAC groups.

Evaluation of the righthand side of Eq. 28 requires use of the Scatchard-Hildebrand equation:

$$\ln(x^s\gamma^s) = \ln\left[\frac{f_S}{f^o_L}\right] = \frac{-\Delta S_f}{R}\left[\frac{T_m}{T} - 1\right] \tag{29}$$

where T is the solution temperature, T_m is the melting point of the solid solute, and ΔS_f is the entropy of fusion of the solute. Equation 29 is based on the identity that

$$\ln\left[f_S/f^o_L\right] = \int_{T_m}^{T} \frac{\Delta H_f}{RT^2}\,dT \tag{30}$$

and on the assumption that ΔH_f, the enthalpy of fusion, is independent of temperature. It also makes use of the identity that

$$\Delta H_f = T_m \Delta S_f \tag{31}$$

The entropy of fusion of the antibiotics is taken in this study to be a constant value of 56.51 J/mol · K (13.5 cal/mol · K). Yalkowsky (1979) reported that the entropy of fusion of many drugs and rigid molecules of intermediate size can be estimated at this value, confirming the observations of Tsonopoulos and Prausnitz (1971).

Using 56.51 J/mol · K for the entropy of fusion permits calculation of γ^s:

$$\gamma^s = \frac{1}{x^s}\exp\left[-6.79\left(\frac{T_m}{T} - 1\right)\right] \tag{32}$$

The only additional information required is T_m.

Experimental data

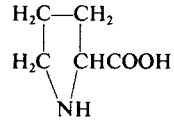
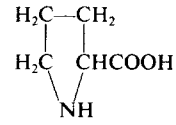
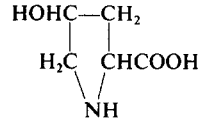
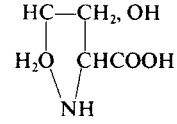
The original measurements of osmotic coefficient data for the amino acids and their conversion to activity coefficients were done by Smith and Smith (1937, 1940a,b) Hutchens et al. (1963), and Ellerton et al. (1964). All these measurements were done at 25°C. The osmotic coefficients were measured by the isopiestic method as suggested by Robinson and Sinclair (1934) and described in Lewis and Randall (1961, p. 265).

A point of subtlety is that the activity coefficients obtained by this method are intended for use with compositions measured in molality and with an infinite dilution standard state, as would be appropriate for nonvolatile solutes. Conversion to the "pure liquid substance" standard state and mole fraction as the concentration measure is discussed by Denbigh (1971). If the symbol γ_A^m is the activity coefficient on the molality basis, then γ_A and γ_A^∞ (the infinite dilution value) are related by:

$$\gamma_A/\gamma_A^\infty = \gamma_A^m/(1 - x_A)$$

Solubilities of amino acids in water at various temperatures were measured by Dalton and Schmidt (1933), Dunn et al. (1933), and Hade (1963).

Table 1. Chemical Structure of Amino Acids

Amino Acid	Chemical Structure	Groups
Glycine	$CH_2(NH_2)COOH$	$CH_2(NH_2)COOH$
Alanine	$CH_3CH(NH_2)COOH$	$CH(NH_2)COOH$, CH_3
Amino-butyric acid	$CH_3CH_2CH(NH_2)COOH$	$CH(NH_2)COOH$, CH_3, CH_2
Valine	$(CH_3)_2CHCH(NH_2)COOH$	$CH(NH_2)COOH$, 2 CH_3, CH
Serine	$HOCH_2CH(NH_2)COOH$	$CH(NH_2)COOH$, CH_2, OH
Threonine	$CH_3CH(OH)CH(NH_2)COOH$	$CH(NH_2)COOH$, CH_3, CH, OH
Proline		
Hydroxy-proline		

Dissociation reaction equilibrium constants pK_1 and pK_2 were measured by Christensen et al. (1967, 1968, 1969) and King (1951).

All the above data for the amino acids have been collected in the *CRC Handbook of Biochemistry and Molecular Biology* (Fasman, 1976). The structures of the amino acids studied in this paper are shown in Table 1.

The chemical structures, molecular weights, and melting temperatures of the antibiotics studied are taken from the *Merck Index* (Windholz, 1976). The chemical formulas and the other data are listed in Table 2.

The antibiotic solubility data employed were measured by Weiss et al. (1957), Andrew and Weiss (1959), and Marsh and Weiss (1967). Many of these data are in the compilation by Tomlinson and Regosz (1985) in the IUPAC *Solubility Data Series*.

Modified UNIFAC parameters

Solution nonidealities have been represented by the modified UNIFAC group contribution model of Larsen et al. (1987). New amino acid and antibiotic groups were defined for this purpose.

New energy interaction parameters have been regressed. These new parameters have not been assumed to be temperature dependent, as is done in modified UNIFAC in general, because the available data are insufficient.

The eight amino acids studied have been represented by two groups (each with two subgroups), a glycine group and a proline group. Alinine, amino-butyric acid, and valine can be thought of as glycine with added methyl (CH, CH_2, or CH_3) groups. Serine and threonine further contain an OH group. Hydroxy-proline is constructed from proline and an OH group.

The six antibiotics studied have been represented by five new large groups. For anisomycin, carbomycin A, griseofulvin, and hygromycin A, the group is the whole molecule. Chloramphenicol is modeled as a large chloramphenicol group and a OH group. Chloramphenicol palmitate contains the chloramphenicol group, the CCOO group, and 15 methyl groups.

The Bondi area and volume parameters for the new groups are listed in Table 3. These quantities, which are required in the modified UNIFAC model, are the sums of the values for the

Table 2. Antibiotics in Study

Antibiotic	Chemical Formula	Melting Temp., K	Molec. Wt.
Anisomycin	$C_{14}H_{19}NO_4$	413.65	265.30
Carbomycin	$C_{42}H_{67}NO_{16}$	487.15	841.97
Chloramphenicol	$C_{11}H_{12}Cl_2N_2O_5$	424.15	323.14
Chloramphenicol palmitate	$C_{27}H_{42}Cl_2N_2O_6$	363.15	561.54
Griseofulvin	$C_{17}H_{17}ClO_6$	493.15	352.77
Hygromycin A	$C_{23}H_{29}NO_{12}$	380.15	511.47

Table 3. Group Size Parameters

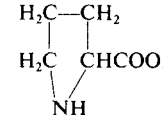
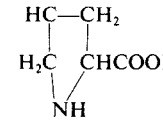
Group Name	Formula	Parameter R	Parameter $(z/2)Q$
Glycine	$CH_2(NH_2)COOH$	2.671	2.914
	$CH(NH_2)COOH$	2.443	2.602
Proline		4.167	3.878
		3.939	3.566
Anisomycin	$C_{14}H_{19}NO_4$	10.1258	7.828
Carbomycin A	$C_{42}H_{67}NO_{16}$	31.6568	25.595
Chloramphenicol	$C_{11}H_{11}Cl_2N_2O_4$	9.8957	7.776
Griseofulvin	$C_{17}H_{17}ClO_6$	11.4753	10.932
Hygromycin A	$C_{23}H_{29}NO_{12}$	19.5745	16.588

Table 4. Group Interaction Parameters

Group k	a_{k,H_2O}	$a_{H_2O,k}$	a_{k,CH_2}	$a_{CH_2,k}$	$a_{k,OH}$	$a_{OH,k}$	$a_{k,ACH}$	$a_{ACH,k}$	$a_{k,CCOO}$	$a_{CCOO,k}$
Glycine	740.9	−13.05	2,281.7	1,916.8	6,769.6	−336.6	—	—	—	—
Proline	346.4	−346.4	—	—	−278.5	123.1	—	—	—	—
Anisomycin	—	—	−582.0	2,152.7	—	—	−491.2	2,691.2	—	—
Carbomycin A	—	—	409.2	−134.4	7,181.9	−204.2	451.6	−183.8	—	—
Chloramphenicol	—	—	129.7	143.8	718.8	−212.7	4,638.0	−137.6	−1,321.8	4,780.0
Griseofulvin	—	—	7,142.0	−92.6	353.5	−145.4	—	—	−4.0	68.4
Hygromycin A	—	—	640.1	−69.4	204.1	75.5	243.9	41.2	76.7	240.5

Unit of measurement: Kelvin

conventional UNIFAC groups that make up the new larger groups.

The optimal a_{ij} parameters obtained in this study are reported in Table 4. The units are Kelvins. These values were regressed using an extended simplex method to minimize appropriate objective functions. The algorithm was adapted for use with modified UNIFAC from the listing in Fredenslund et al. (1977). In the case of the amino acids, the objective function was;

$$\text{O.F.} = \Sigma\,[\ln(\gamma_i/\gamma_i^\infty)_{exp} - \ln(\gamma_i/\gamma_i^\infty)_{UNIFAC}]^2$$

For the antibiotics,

$$\text{O.F.} = \Sigma\,[\ln(\gamma_i^s) - \ln(\gamma_i^s)_{UNIFAC}]^2$$

where the γ_i^s values were calculated from saturation data using Eq. 32.

Results

Amino acids

Seventeen data points, γ/γ^∞, for glycine and alanine in neutral water were used to deduce the glycine-H_2O and glycine-CH2 modified UNIFAC a_{ij} values.

Fifteen data for proline in water were used in fitting the proline-H_2O energy interaction parameters. The proline-OH a_{ij} parameters were then obtained from the hydroxy-proline activity coefficient data.

From the parameters obtained for the glycine group, it was possible to predict γ/γ^∞ for solutions of amino-butyric acid, valine, and threonine. These substances were not used in the parameter regressions.

Table 5. Overall Representation of Experimental γ Data for Amino Acids with Modified UNIFAC

Amino Acid	No. Data Points	RMS Error in Prediction %	Predicted γ^∞ at 298.15 K	No. Parameters Regressed
Glycine	10	4.20	9.20	4
Alanine	7	8.97	6.43	4
Amino-butyric acid	7	17.34	10.00	0
Valine	3	12.15	17.90	0
Serine	11	3.32	4.88	2
Threonine	7	13.78	5.79	0
Proline	15	3.01	0.08	2
Hydroxy-proline	7	0.06	0.18	2
Overall	67	8.52		10

Table 5 presents the RMS errors in the activity coefficient ratios for the eight amino acids. The infinite dilution activity coefficients that arise from the modified UNIFAC correlations are also given in the table. The RMS errors are not large and the γ^∞ values are reasonable. Of course, γ^∞ cannot be known definitely from the data used and the values in Table 5 must be regarded only as arising as a consequence of the correlation effort.

Figure 1 shows the variation with mole fraction of the activity coefficient ratios for the five amino acids used in the regression of parameters. For four of these, the results are clearly acceptable. However, the curve for alanine does not pass through the experimental data (open squares). In fact, the slope of the correlation is incorrect, with γ/γ^∞ decreasing with increasing mole fraction while the data show the opposite trend. Apparently, the glycine-CH_2 parameters obtained in the regression do not permit correct representation of this important qualitative behavior.

Figure 2 contains the predictive results. The curves for amino-butyric acid and valine show the same qualitative failure as in the alanine results. These two amino acids are obtained from alanine by the successive addition of methyl groups, so the difficulties are perhaps not surprising. In contrast, the activity coefficient ratio for threonine decreases with increasing mole fraction both in the data and in the UNIFAC model.

Equation 24 has been used to complete the solubility model for three of the amino acids, dl-alanine, glycine, and dl-valine. Solubility data for these acids in neutral water at various temperatures were available. The UNIFAC model cannot differentiate between optical isomers, therefore the data for the mixed forms of alanine and valine were used. At the experimental mole

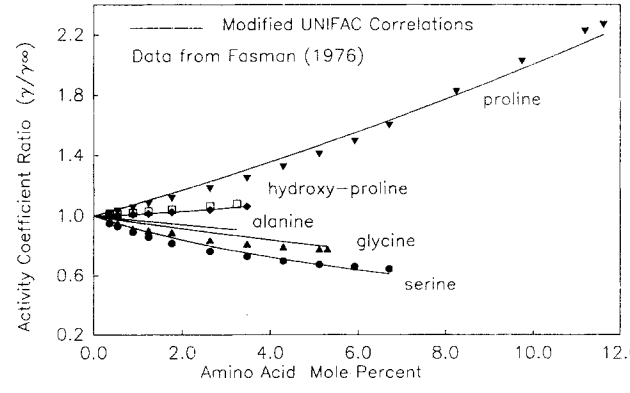


Figure 1. Modified UNIFAC correlation of activity coefficients for five amino acids in neutral water.
□Alanine data

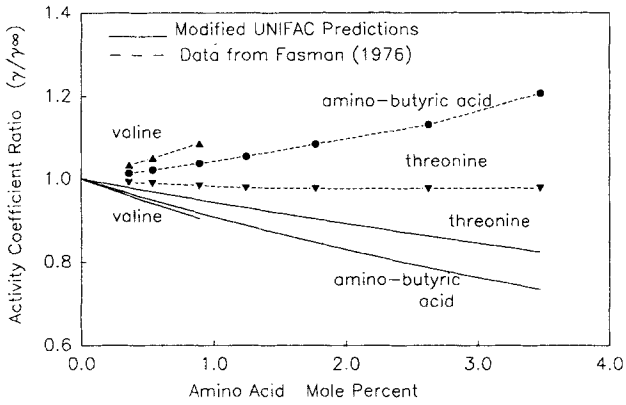


Figure 2. Modified UNIFAC prediction of activity coefficients for three amino acids in neutral water.

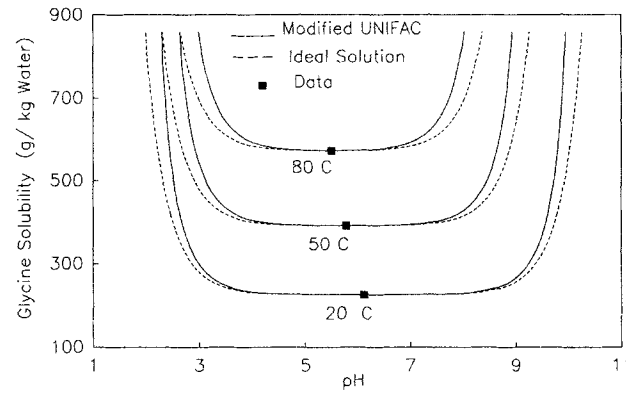


Figure 3. Predicted solubility of glycine in water vs. pH at three temperatures.
Parameters adjusted to match minimum solubility points

fractions, the pH of the solutions was calculated from the modified UNIFAC model. These results were inserted into Eq. 24 to obtain numerical values of $\ln(f_{A_S}/f^{\circ}_{A_{\pm L}})$. The variations with temperature of the pK values and the activity coefficients were included in this calculation. Coefficients A' and B' were regressed from the values of $\ln(f_{A_S}/f^{\circ}_{A_{\pm L}})$ at several temperatures and the results are given in Table 6.

Equation 24 can now be used to predict the effect of pH and temperature on the solubilities of the three amino acids. The results of these predictions for glycine are shown in Figure 3. The curves at each of three temperatures show the expected minimum solubility at the isoelectric pH values. Because of the method used in determining parameters, the isoelectric solubilities are predicted exactly. The solubilities at pH values on either side of the isoelectric point correspond to mole fractions greater than any at which activity coefficients could be measured and therefore imply extrapolation of the activity coefficient model into composition ranges outside that used in fitting parameters.

The calculated solubility of glycine is very sensitive to the activity coefficients. To demonstrate this point, computations have been done with an "ideal solution" version of Eq. 24 in which the minimum solubility is matched exactly but the activity coefficient is presumed constant. The resulting curves are also shown in Figure 3. The calculated ideal solution solubility is always less than the solubility calculated assuming variable activity coefficients, and the differences are large at pH values removed from the isoelectric point.

Antibiotics

Group energy interaction parameters for the antibiotic groups are presented in Table 4. Thirty-two a_{ij} values have been found from 32 measured solubilities of the six antibiotics in a total of ten different solvents. The solubilities were at 28°C except for griseofulvin at 21°C.

Saturation data were available for anisomycin in cyclohexane, iso-octane, benzene, and toluene. These four measurements were converted to four γ^s-values. In separate regressions, the anisomycin-methyl and anisomycin-ACH energy interaction parameters were found.

The other antibiotic groups were treated similarly. The four a_{ij} values between carbomycin A and the CH_2 and OH groups were regressed from solubilities in three alkanols and cyclohexane. The carbomycin-ACH parameters were found from solubilities in benzene and toluene. The eight required parameters for the chloramphenicol group were regressed simultaneously from eight solubility data for chloramphenicol and chloramphenicol palmitate in cyclohexane, iso-octane, two acetates, one alkanol, benzene, toluene, and benzyl alcohol. Six griseofulvin parameters were found in two separate regressions using six solubility measurements. Three separate regressions were used to evaluate the eight hygromycin A parameters.

In each case, the number of parameters to be found just equaled the number of data. In spite of this, it was often impossible to match the data exactly; the regressions often resulted in a minimum (but nonzero) value for the objective function. This unexpected result was examined with some care and verified by examining the values of the parameters for which each of the equations could be satisfied separately. Perhaps the explanation for this result is in the very large values for the group size parameters and the resulting large contribution to the activity coefficients of the UNIFAC combinatorial terms. The energy interaction parameters affect only part of the activity coefficient in the UNIFAC model and, apparently, this part did not contain sufficient flexibility to compensate for large effects elsewhere in the model.

The saturation activity coefficients calculated from the solubility data and the values obtained as a result of the regressions are summarized in Table 7. The range of activity coefficient values is from 3.4 for griseofulvin in ethyl acetate to 32,500 for hygromycin A in benzene. The maximum deviation is for carbomycin A in cyclohexane (55%). Large errors, in general, are associated with very low solubilities.

The four anisomycin activity coefficients were matched exactly. However, an examination of the activity in subsaturated solutions in cyclohexane, benzene, and toluene shows that these

Table 6. $(f^{\circ}_{A_S}/f^{\circ}_{A_{\pm L}})$, T Correlation Constants*

Amino Acid	A'	B' K
dl-Alanine	2.060	1,110.971
Glycine	2.042	885.689
dl-Valine	2.305	1,217.946

*$\ln(f^{\circ}_{A_S}/f^{\circ}_{A_{\pm L}}) = A - \frac{B'}{T}$

Table 7. Antibiotic Activity Coefficients at Saturation

Solvent	Values from Solubility Data (Model Value) % Deviation					
	Anisomycin	Carbomycin A	Chloramphenicol	Chloramphenicol Palmitate	Griseofulvin	Hygromycin A
Cyclohexane	1,658	267	—	3,831	1,317	21,530
	(1,658) 0%	(414) 55%	—	(4,830) 26.1%	(1,315) −0.2%	(21,510) −0.1%
Isooctane	12,700	1,181	—	9,881	—	—
	(12,700) 0%	(764) − 35%	—	(7,877) −20.2%	—	—
Benzene	70.1	7.73	878	147	—	32,500
	(70.1) 0%	(8.78) 13.5%	(868) −1%	(148) 0.6%	—	(32,500) 0%
Toluene	130	31.6	1,307	96.7	—	—
	(130) 0%	(26.7) −15%	(1,316) 0.7%	(98.4) 1.8%	—	—
Ethanol	—	—	—	—	11.2	91.7
	—	—	—	—	(14.4) 28.7%	(98.1) 6.9%
Isopropanol	—	35.8	—	—	35.4	102.4
	—	(36.3) 1.4%	—	—	(22.5) −36.6%	(90.2) −11.9%
Isoamyl alcohol	—	116.9	10.8	—	31.4	82.4
	—	(114.8) −1.8%	(10.9) 0.9%	—	(38.6) 22.8%	(87.6) 6.3%
Benzyl alcohol	—	—	16.8	—	—	92.1
	—	—	(16.3) −2.6%	—	—	(92.1) 0%
Ethyl acetate	—	—	—	—	3.43	2,933
	—	—	—	—	(3.43) 0%	(2,933) 0%
Isoamyl acetate	—	—	—	—	8.3	1,449
	—	—	—	—	(8.3) 0%	(1,440) 0%

liquids will be unstable at saturation and the model has to be regarded as a failure. This unacceptable model behavior can be eliminated by accepting smaller values for the anisomycin surface area parameter; however, the available data were not sufficient to justify any further manipulation of parameters.

With caution, the group interaction parameters can be used to predict solubilities in solvents other than those for which data are available and in mixed solvents. The parameter table is sufficient, for some of the antibiotic groups, to consider the broad range of alkanes, aromatics, alcohols, and esters.

Antibiotics are often produced from solution as crystalline precipitates by adding solvents in which the solubility is low. Figure 4 shows the calculated solubility of carbomycin A in mixtures of cyclohexane and toluene at 28°C (the temperature at which solubilities in the pure solvents were measured). Comparisons with data are possible at the endpoints of the curve; on these axes, the errors do not appear especially significant. The solubility in the mixtures does not vary linearly with the cyclohexane mole fraction and definitely reflects a contribution from the activity coefficient model.

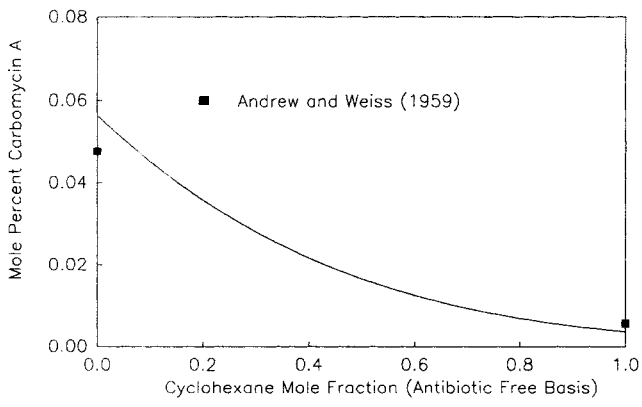


Figure 4. Predicted solubility of carbomycin A in mixtures of toluene and cyclohexane.

An alternative way (closer to the processing applications) of looking at the solubilities in mixed solvents is presented in Figure 5. The figure shows the calculated percent recovery of carbomycin A that could be expected on diluting a saturated solution in toluene with several solvents. According to these calculations, nearly 90% of the dissolved antiobiotic could be recovered as crystalline solid by diluting with *n*-hexadecane on a one-to-one basis. The curves in this figure are typical of the predictive calculations possible.

Comment

As shown in Figure 3, it is relatively easy to model exactly the solubilities of the amino acids at their isoelectric points. Only the parameters in Table 6 need be fitted since other essential physicochemical data are well known. Away from this minimum solubility point, however, an activity coefficient model will be essential if the highly sensitive behavior is to be fit with any precision.

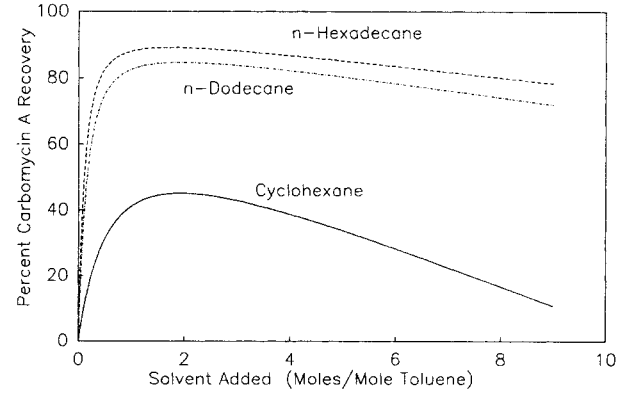


Figure 5. Predicted recovery of carbomycin A from saturated solutions in toluene on dilution with selected hydrocarbons.

Unfortunately, there were no data found for the solubilities of the amino acids we considered with varying solution pH. It was therefore necessary to use very limited and rather old data as the basis of a correlation effort and then extrapolate into the concentration ranges of interest in a predictive effort.

The modified UNIFAC model, as we used it, proved to be a flawed tool in predicting even the limited activity coefficient data in the series of glycine, alanine, amino-butyric acid, and valine. It could only be used with caution, and the solubility behavior vs. pH for the higher amino acids in this series will, most probably, be inaccurate.

The nature of the activity coefficient data available in Fasman (1976) also deserves scrutiny. These data were obtained from measurements of water activity by integrating the Gibbs-Duhem equation to determine the amino acid activity indirectly. The procedure ignores the detailed ionic behavior of the amino acid solute and leaves no alternative to the assumption (which we have used) that all the various amino acid ionic species have the same activity coefficient in solution.

Our antibiotic solubility model is also based on a very limited data base. The predictions in Figures 4 and 5 certainly would require some experimental verification before use in a design. Fortunately, experiments of the kind indicated in the figures could be carried out without too much difficulty for the mixed solvents of interest in a particular antibiotic recovery process.

Our purpose was to demonstrate proposed methods for dealing with the solubilities of two important classes of biological compounds. A model for the behavior of a specific substance would certainly have to be evaluated by comparison with additional data.

Acknowledgment

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

Notation

a = group interaction parameter
A = amino acid
A', B' = coefficients, Eq. 25
f = fugacity
H = enthalpy
K = equilibrium constant
m = molality
n_o = net amino acid molality
O.F. = objective function
$pK = -\log_{10} K$
Q = UNIFAC surface area parameter
R = UNIFAC volume parameter; gas constant
S = entropy
T = temperature
x = mole fraction
x = mole fraction
x_A^o = net amino acid mole fraction
z = UNIFAC contact number parameter

Greek letters

ϵ = reaction extent
γ = activity coefficient

Subscripts

A = amino acid
exp = experimental value
D = dissociation
f = fusion
i,j = group, substance, or reaction index
L = liquid
m = melting
S = solid
w = water
$\pm$ = amino acid zwitterion

Superscripts

m = molality basis
q = standard state
s = saturation
∞ = infinite dilution

Literature Cited

Andrew, M. L., and P. J. Weiss, "Solubility of Antibiotics in Twenty-four Solvents. II," *Antibiot. Chemotherapy,* **9,** 277 (1959).

Bondi, A., *Physical Properties of Molecular Crystals, Liquids and Glasses,* Wiley, New York (1968).

Christensen, J. J., R. M. Izatt, and L. D. Hansen, "Thermodynamics of Proton Ionization in Dilute Aqueous Solution. VII: ΔH^o and ΔS^o Values for Proton Ionization from Carboxylic Acids at 25°," *J. Amer. Chem. Soc.,* **89,** 213 (1967).

Christensen, J. J., J. L. Oscarson, and R. M. Izatt, "Thermodynamics of Proton Ionization in Dilute Aqueous Solution. X: $\Delta G^o(pK)$, ΔH^o and ΔS^o for Proton Ionization from Several Monosubstituted Carboxylic Acids at 10, 25, and 40°," *J. Amer. Chem. Soc.,* **90,** 5949 (1968).

Christensen, J. J., R. M. Izatt, D. P. Wrathall, and L. D. Hansen, "Thermodynamics of Proton Ionization in Dilute Aqueous Solution. XI: pK, ΔH^o, and ΔS^o Values for Proton Ionization from Protonated Amines at 25°," *J. Chem. Soc., A,* No. 8, 1212 (1969).

Dalton, J. B., and C. L. A. Schmidt, "The Solubilities of Certain Amino Acids in Water, the Densities of their Solutions at 25°, and the Calculated Heat of Solution and Partial Molal Volumes," *J. Biol. Chem.,* **103,** 549 (1933).

Denbigh, K., *The Principles of Chemical Equilibrium,* 3d Ed., Cambridge Univ. Press, London (1971).

Dunn, M. S., F. J. Ross, and L. S. Read, "The Solubility of the Amino Acids in Water," *J. Biol. Chem.,* **103,** 579 (1933).

Ellerton, H. D., G. Reinfelds, D. E. Mulcahy, and P. J. Dunlop, "Activity, Density, and Relative Viscosity Data for Several Amino Acids, Lactamide, and Raffinose in Aqueous Solution at 25°," *J. Phys. Chem.,* **68,** 398 (1964).

Fasman, G. D., ed., *Handbook of Biochemistry and Molecular Biology,* 3d Ed., *Physical and Chemical Data,* vol. 1, CRC Press, Cleveland (1976).

Fredenslund, Aa., J. Gmehling, and P. Rasmussen, *Vapour Liquid Equilibria Using UNIFAC,* Elsevier, New York (1977).

Greenstein, J. P., and M. Winitz, *Chemistry of the Amino Acids,* vol. 1, Wiley, New York (1961).

Hade, E. P. K., "Solubilities of Naturally Occurring L-Amino Acids in Water at Various Temperatures," Ph. D. Thesis, U. Chicago (1963).

Hutchens, J. O., K. M. Figlio, and S. M. Granito, "Isopiestic Comparison Method for Activities—Activities of L-Serine and L-Arginine-HCl," *J. Biol. Chem.,* **238,** 1419 (1963).

King, E. J., "The Ionization Constants of Glycine and the Effect of Sodium Chloride upon Its Second Ionization," *J. Amer. Chem. Soc.,* **73,**155 (1951).

Larsen, B. L., P. Rasmussen, and Aa. Fredenslund, "A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heat of Mixing," *Ind. Eng. Chem. Res.,* **26,** 2274 (1987).

Lewis, G. N., and M. Randall, *Thermodynamics,* 2d ed. (rev. by K. S. Pitzer and L. Brewer), McGraw-Hill, New York (1961).

Marsh, J. R., and P. J. Weiss, "Solubility of Antibiotics in Twenty-six Solvents. III," *J. Assn. Official Analyt. Chemists,* **50,** 457 (1967).

Nass, K. K., "Representation of the Solubility Behavior of Amino Acids in Water," *AIChE J.,* **34,** 1257 (1988).

Orella, C. J., and D. J. Kirwan, "Correlation of the Solubility of Amino Acids and β-Lactum Antibiotics in Aqueous/Organic Solvent Systems," paper presented at AIChE, Ann. Meet., New York (1987).

Robinson, F. A., and D. A. Sinclair, "The Activity Coefficients of the Alkali Chlorides and Lithium Iodide in Aqueous Solution from Vapor Pressure Measurements," *J. Amer. Chem. Soc.,* **56,** 1830 (1934).

Smith, P. K., and E. R. B. Smith, "Thermodynamic Properties of Solutions of Amino Acids and Related Substances. II: The Activity of Aliphatic Amino Acids in Aqueous Solution at 25°," *J. Biol. Chem.*, **121,** 607 (1937).

———, "Thermodynamic Properties of Solutions of Amino Acids and Related Substances. V: The Activities of Some Hydroxy- and *n*-Methylamino Acids and Proline in Aqueous Solution at 25°," *J. Biol. Chem.*, **132,** 57 (1940a).

———, "Thermodynamic Properties of Solutions of Amino Acids and Related Substances. VI: The Activities of Some Peptides in Aqueous Solution at 25°," *J. Biol. Chem.*, **135,** 273 (1940b).

Tomlinson, E., and A. Regosz, eds., *Solubility Data Series*, vols. 16/17, *Antibiotics: I, β-Lactum Antibiotics*, Pergamon, Oxford (1985).

Tsonopoulos, C., and J. M. Prausnitz, "Activity Coefficients of Aromatic Solutes in Dilute Aqueous Solutions," *Ind. Eng. Chem. Fundam.*, **10,** 593 (1971).

Weiss, P. J., M. L. Andrew, and W. W. Wright, "Solubility of Antibiotics in Twenty-four Solvents," *Antibiot. Chemotherapy*, **7,** 374 (1957).

Windholz, M., ed., *The Merck Index*, Merck & Co., Rahway, NJ (1976).

Yalkowsky, S. H., "Estimation of Entropies of Fusion of Organic Compounds," *Ind. Eng. Chem. Fundam.*, **18,** 108 (1979).

Manuscript received Aug. 22, 1989, and revision received Dec. 7, 1989.