Correlation and Prediction of Crystal Solubility and Purity

Amyn S. Teja, Jeffrey C. Givand, and Ronald W. Rousseau
School of Chemical Engineering, Georgia Institute of Technology, Atlanta, GA 30332

A thermodynamic model is developed for the correlation and prediction of crystal solubility and purity when lattice substitution by isomorphic or near-isomorphic impurities is the dominant mechanism that leads to impure crystals. The model is based on earlier observation that the purity of amino acid crystals obtained from aqueous solutions is related to the relative solubilities of the product and impurity in the solvent. A simple relationship is developed that allows the prediction of the purity of crystals obtained from dilute solutions containing both product and impurity species.

Introduction

Crystallization of a solute from a solution can be induced by cooling, addition of a nonsolvent, solvent evaporation, or a number of other methods. It is common to find that the solution (mother liquor) is contaminated with impurities, and these impurities are often isomorphic (similar in molecular functionality and size) or near-isomorphic with the primary species. Byproducts from chemical or biochemical reactions are examples of such impurities, and their removal from the desired product is of importance in many applications.

The contamination of a crystalline product by a near-isomorphic impurity is controlled by a number of phenomena, including the thermodynamics of the solid–liquid system. Here we consider the dominant mechanism of impurity capture to be substitution of the impurity into the lattice of the primary product. In addition, we assume that the substituted lattice behaves as a solid solution, and that equilibrium between the solid and liquid phases represents a limiting condition that can be approached in a crystallization process, although it is likely that equilibrium is never truly achieved in practice. Accordingly, the present work addresses the development of a thermodynamic framework that can be used to predict, at least qualitatively, the effect of the solvent and its characteristics on crystal purity. The development of the framework involves correlations of solubility that have utility in their own right, and provide an additional useful aspect of this work.

We demonstrate the application of the thermodynamic framework to the case of amino acid crystallization, both because amino acids are important in food and pharmaceutical applications, and because these compounds serve as models for other biologically important isomorphic molecules. As an example, L-isoleucine (L-ILE) is widely used as an active ingredient in pharmaceutical preparations and as a food additive, and its isolation from fermentation broths that also contain small amounts of other amino acids, such as L-leucine (L-LEU) and L-valine (L-VAL), is typical of many bioseparations.

The solubilities of amino acids have been correlated by a number of researchers, including Chen et al. (1989), Orella and Kirwan (1989, 1991), Gupta and Heidemann (1990), Peres and Macedo (1994), and Pinho et al. (1994). In particular, Orella and Kirwan (1989, 1991) and Gupta and Heidemann (1990) proposed activity-coefficient approaches to estimate the solubility of amino acids in mixed-solvent systems. However, few attempts have been made to relate the solubility to crystal purity. A notable exception is the work of Rosenberger and Riveros (1974), who derived a relationship between the segregation coefficients for the crystallization of alkali halides from pure water to the heats of dissolution and solubilities of the halides in water. We have previously (Givand et al., 1999, 2001, 2002) presented experimental data showing that the purity of amino acid crystals obtained from aqueous solutions (with or without cosolvents) is related to the relative solubilities of the pure species in the solvent. This article presents a thermodynamic framework that quantitatively relates the relative solubilities to the purity of crystals. The model should prove useful in identifying cosolvents and/or additives that, when added to dilute aqueous solutions of two or more isomorphic compounds, would yield product crystals of enhanced purity.

Correspondence concerning this article should be addressed to A. S. Teja.

Thermodynamic Development

Solubility model (binary or pseudobinary systems)

In the development of the thermodynamic framework, we first consider the case of a product species that crystallizes in pure form from a solution. We denote the solvent (whether pure or mixed) as component 1 and the solute (or product) as component 2. For equilibrium, we can write

$$f_2^S = x_2^o \gamma_2^o f_2^*, \tag{1}$$

where f_2^S is the fugacity of 2 in the (pure) solid phase, x_2^o is the solubility of pure 2 in the solvent, γ_2^o is the liquid-phase activity coefficient of 2, and f_2^* is the standard-state fugacity to which γ_2^o refers. As discussed by Prausnitz et al. (1999), it is convenient to define the standard-state fugacity as that of the pure subcooled liquid at the temperature of the solution. Use of a thermodynamic cycle then leads to

$$\ln\left(x_2^o \gamma_2^o\right) = \ln\left(\frac{f_2^S}{f_2^*}\right) = \frac{\Delta h_2^{\text{fus}}}{RT_{t2}}\left(1 - \frac{T_{t2}}{T}\right) + \frac{\Delta C_{P2}}{RT_{t2}}\left(1 - \frac{T_{t2}}{T}\right) - \frac{\Delta C_{P2}}{R}\ln\left(\frac{T_{t2}}{T}\right), \tag{2}$$

where Δh_2^{fus} is the enthalpy change upon melting of the solute at its triple (or melting) -point temperature, T_{t2}, and ΔC_{P2} is the difference between the heat capacities of the solute in the liquid and solid phases. The righthand side of Eq. 2 therefore contains the physical properties of the solute only, with the influence of the solvent on the solute solubility being accounted for by the activity coefficient γ_2^o on the lefthand side of the equation. The righthand side therefore can be evaluated if measured or estimated values of the physical properties (Δh_2^{fus}, T_{t2}, and ΔC_{P2}) of the solute are available. Solubility data can then be correlated using an activity-coefficient model to obtain γ_2^o. Unfortunately, most amino acids decompose at temperatures below their melting points, and physical property data required in Eq. 2 are not readily available. We may, however, neglect the ΔC_P terms, as these are often small compared with the Δh_2^{fus} terms. Moreover, the activity coefficient in Eq. 2 can be replaced by its value at infinite dilution when the systems under study are dilute (as is the case for the amino acid systems examined in this work). Finally, a simple relationship can be adopted for the infinite-dilution activity coefficient as follows

$$\ln \gamma_2^\infty = A_2 + \frac{B_2}{T}. \tag{3}$$

This leads to the following simple expression for the solubility in such systems

$$\ln x_2^o = A'_2 + \frac{B'_2}{T}, \tag{4}$$

with

$$A'_2 = \frac{\Delta h_2^{\text{fus}}}{RT_t} - A_2 \tag{4a}$$

$$B'_2 = -\frac{\Delta h_2^{\text{fus}}}{R} - B_2. \tag{4b}$$

In principle, Eq. 4 can be used to correlate solubility data for any solute in any solvent, provided the solution is dilute. Constants obtained by fitting solubility data for a specific solute (e.g., the amino acids in the present study) in different solvents provide an indication of the effect of solution nonidealities on the solubility of the solute (since the first terms in both Eqs. 4a and 4b have fixed values for that solute).

Model for crystal purity (ternary or pseudoternary systems)

Extension of the model to the case of a solid solution of two solutes (primary species plus an impurity) in equilibrium with a dilute solution containing the two solutes can be accomplished as follows. For $i = 2$ or 3 (the solvent is component 1)

$$z_i \Gamma_i f_i^S = x_i \gamma_i f_i^*, \tag{5}$$

where z_i and Γ_i are the mole fraction and activity coefficient of the solute i (product or impurity) in the solid phase, and x_i and γ_i are the corresponding quantities in the liquid phase. Eliminating the pure-component fugacity ratio using Eq. 1

$$\frac{\gamma_i x_i}{z_i \Gamma_i} = \gamma_i^o x_i^o. \tag{6}$$

Equation 6 can be written for product 2 as well as for impurity 3, and the two equations can be combined with the fact that $z_2 + z_3 = 1$ in the crystal, to give

$$z_2 = \frac{\left(\dfrac{x_2}{x_3}\right)\left(\dfrac{x_3^o}{x_2^o}\right)\left(\dfrac{\gamma_3^o \Gamma_3/\gamma_3}{\gamma_2^o \Gamma_2/\gamma_2}\right)}{1 + \left(\dfrac{x_2}{x_3}\right)\left(\dfrac{x_3^o}{x_2^o}\right)\left(\dfrac{\gamma_3^o \Gamma_3/\gamma_3}{\gamma_2^o \Gamma_2/\gamma_2}\right)}. \tag{7}$$

All quantities on the righthand side of this relationship are known or can be obtained from solubility correlations, as discussed below. The ratio of x_3 to x_2 represents the molar impurity ratio in solution at the equilibrium temperature. It can be determined by sampling the solution (or, if the amount of crystals formed is small, it can be obtained from the impurity ratio in the mother liquor). The ratio of x_3^o to x_2^o is obtained from the solubilities of the pure amino acids (2 and 3) in the given solvent. The only unknown in Eq. 6 is, therefore, the term containing activity coefficients. This can be estimated if we have solid–liquid equilibrium data in one solvent (for example, water), as discussed below.

As an alternative to the approach just described, Eq. 6 for the product and impurity can be combined to give

$$\beta_{32} = \left(\frac{K_3}{K_2}\right) = \frac{z_3/x_3}{z_2/x_2} = \left(\frac{\gamma_2^o \Gamma_2/\gamma_2}{\gamma_3^o \Gamma_3/\gamma_3}\right)\left(\frac{x_2^o}{x_3^o}\right) \tag{8}$$

or

$$\beta_{32}\,\alpha_{32}^o = \left(\frac{\gamma_2^o \Gamma_2/\gamma_2}{\gamma_3^o \Gamma_3/\gamma_3}\right), \tag{8a}$$

where $K_i = z_i/x_i$ is the distribution coefficient of solute i between the solid (crystal) and liquid (mother liquor) phases, and x_i^o is the solubility of the pure solute i (product or impurity) in the solvent at the system temperature. The distribution of the impurity relative to the product, β_{32}, therefore depends on the relative solubility $\alpha_{32}^o (= x_3^o/x_2^o)$ of the pure species. Furthermore, the product $\beta_{32}\alpha_{32}^o$ provides a measure of the nonideality of the system arising from both the liquid and solid phases.

Results

Solubility (binary or pseudobinary systems)

The solubilities of L-isoleucine, L-leucine, and L-valine in water between 293 and 333 K were obtained from the literature (see Givand et al., 1999, 2001, 2002) and correlated using Eq. 4. The results of the correlations are presented in Figure 1 and Table 1. Table 1 shows that the data were well-correlated with the model yielding R-values close to one. The good fits to the data suggest that the infinite-dilution assumption is not unreasonable for the aqueous amino acid solutions examined. Equation 4 was also used to correlate the solubilities of L-isoleucine and L-leucine in mixed aqueous solvents. Table 2 summarizes the A' and B' parameters (and R-values) obtained from these correlations. Once again, R-values were close to unity, indicating good fit of the data. Typical data and correlation results are plotted in Figure 2.

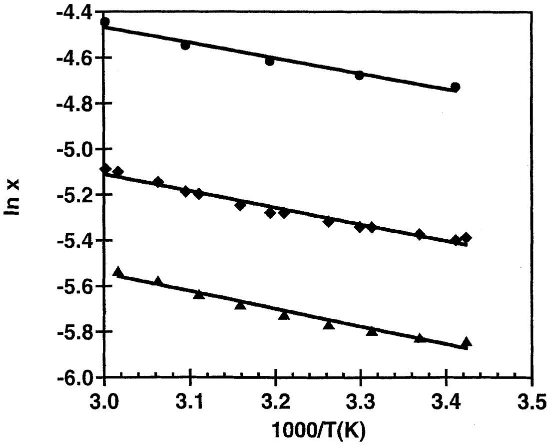

Figure 1. Experimental and correlated solubilities of *L*-leucine (▲), *L*-isoleucine (♦), and *L*-valine (●) in water.

Solid–liquid equilibrium (ternary or pseudoternary systems)

The activity coefficient ratio $\gamma_2^o\Gamma_2/\gamma_2/\gamma_3^o\Gamma_3/\gamma_3$ cannot be set to 1 because of the presumed energetic differences in the lattice substitution of the amino acids, and the nonideality of the solutions. Data from experiments in which the amino acids were crystallized from pure water at 297 K were therefore used to obtain this ratio. Measured crystal purities (Givand et al., 2001) were correlated using Eq. 7, with the activity coefficient ratio treated as a parameter.

Figure 3 shows a plot of the mole fraction of the impurity (L-leucine) in L-isoleucine crystals obtained from an aqueous solution containing both L-leucine and L-isoleucine. Note that if the product (L-isoleucine) is crystallized from and is in equilibrium with a solution in which the impurity (L-leucine) comprises 10 mol % of the solution on a solvent-free basis, then the crystals will contain ~12.5 mol % L-leucine. Hence the impurity concentration is higher in the crystals than in the starting mother liquor—a situation that is obviously undesirable. The dashed line in the figure would be obtained if the impurity concentration in the crystals is ex-

Table 1. Parameters of Eq. 8 and *R*-Values for *L*-Leucine, *L*-Isoleucine, and *L*-Valine Solubilities in Water between 293 and 333 K

Model Parameter	Isoleucine	Leucine	Valine
A'	−2.923	−3.223	−2.453
B'	−729.4	−773.5	−671.6
R-value	0.985	0.985	0.985

Table 2. Parameters of Eq. 8 and *R*-Values for *L*-Leucine and *L*-Isoleucine Solubilities in Various Aqueous Solvent Mixtures Between 293 and 333K

	Isoleucine			Leucine		
Aqueous Mixture	A'	B'	R	A'	B'	R
$CaCl_2$ ($I = 1.27$ mol·kg^{-1})	−2.741	−755	0.993	−2.924	−829	0.991
$(NH_4)_2SO_4$ ($I = 0.75$)	−2.752	−788	0.986	−3.041	−832	0.989
$(NH_4)_2SO_4$ ($I = 1.62$)	−2.705	−838	0.989	−2.950	−898	0.986
$(NH_4)_2SO_4$ ($I = 4.01$)	−2.554	−982	0.985	−2.786	−1,045	0.983
8.0 mol % t-butanol	1.152	−2,161	0.999	1.133	−2,298	1.000
15.0 mol % t-butanol	0.992	−2,202	0.989	1.494	−2,496	1.000
20.4 mol % ethylene glycol	−0.840	−1,664	0.999	−1.083	−1,711	0.999
10.1 mol % DMSO	−1.201	−1,619	0.994	−1.478	−1,653	0.997
10.1 mol % DMSO + $CaCl_2$ ($I = 1.23$)	−0.405	−1,781	0.996	−0.794	−1,777	0.996
20.2 mol % DMSO + $CaCl_2$ ($I = 1.23$)	0.836	−2,402	0.999	0.355	−2,355	0.996
20.5 mol % DMSO + $CaCl_2$ ($I = 3.02$)	1.318	−2,417	0.999	0.929	−2,390	0.999

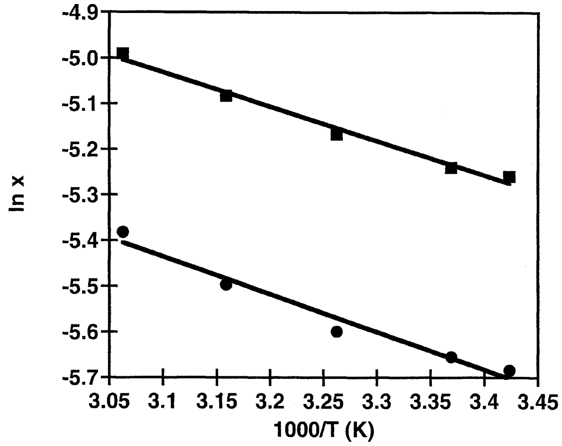

Figure 2. Experimental and correlated solubilities of *L*-leucine (■) and *L*-isoleucine (●) in aqueous $CaCl_2$ solutions of ionic strength $I = 1.27$ mol·kg^{-1}.

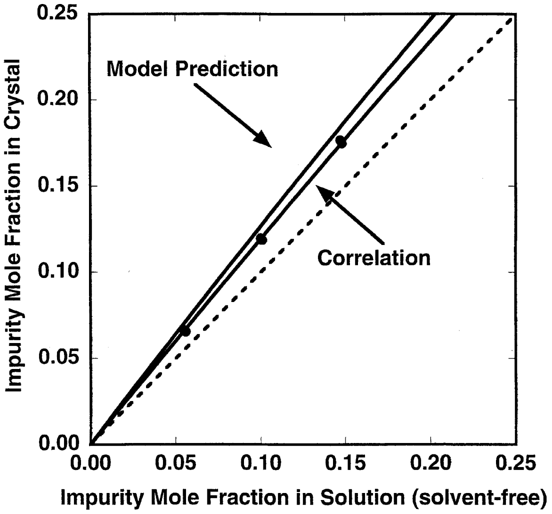

Figure 4. Impurity (*L*-leucine) content of *L*-isoleucine crystals obtained from solutions of *L*-leucine and *L*-isoleucine in aqueous $CaCl_2$ solutions (of ionic strength $I = 1.6$ mol·kg^{-1}) at 297 K.

Predictions are given by the solid line.

actly the same as the (solvent-free) impurity concentration in the mother liquor. The solid curve in Figure 3 shows a fit of the purity data using a constant value for the activity coefficient ratio $\gamma_2^o\Gamma_2/\gamma_2/\gamma_3^o\Gamma_3/\gamma_3$. As can be seen in this figure, the data are fit very well with a constant value of this ratio. The reasons why this might be so are discussed in the following section.

Figures 4 and 5 show predicted and experimental crystallization data for *L*-isoleucine or *L*-leucine from aqueous solutions containing cosolvents and/or additives. The predicted curves were obtained using Eq. 7 and activity-coefficient ratios from data for *L*-isoleucine or *L*-leucine crystallized from their solutions in water. Enhancement of *L*-isoleucine crystal purity is predicted by the model, as is deterioration in *L*-leucine crystal purity, in agreement with the data. The dashed lines in the figures would be obtained if the impurity concentration in the crystals was exactly the same as the (solvent-free) impurity concentration in the mother liquor (that is, if there was no enhancement or deterioration in the concentration of the impurity). These predictions lead to the remarkable conclusion that the activity coefficient ratio

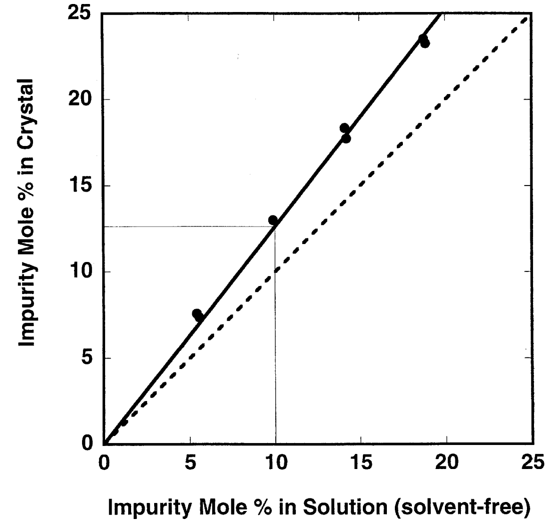

Figure 3. Impurity (*L*-leucine) content of *L*-isoleucine crystals obtained from solutions of *L*-leucine and *L*-isoleucine in water at 297 K.

The solid line represents the correlation (Eq. 7), and the dashed line represents impurity concentrations in the crystals that are exactly the same as the (solvent-free) impurity concentrations in the mother liquor.

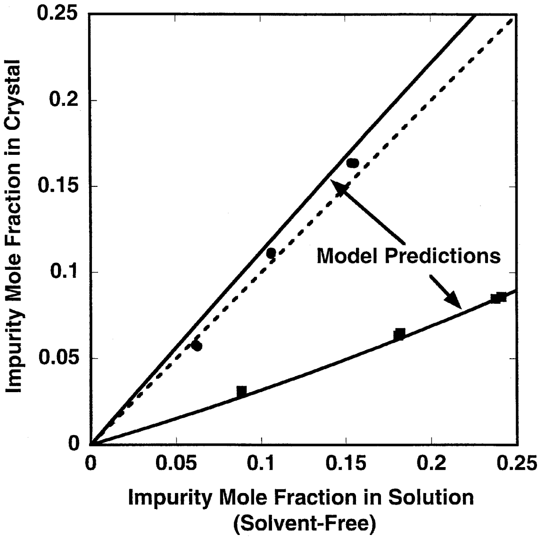

Figure 5. Impurity (*L*-leucine) content of *L*-isoleucine crystals obtained from solutions of *L*-leucine and *L*-isoleucine in aqueous DMSO (10 mol %) + $CaCl_2$ ($I = 3.1$ mol·kg^{-1}) at 297 K.

Predictions are given by the solid line.

Table 3. Relative Distributions and Relative Solubilities when *L*-Isoleucine (component 2) is Crystallized from Aqueous Systems in the Presence of *L*-Leucine (component 3) at 297 K

Aqueous Systems	β_{32}	α_{32}^o	$\beta_{32}\alpha_{32}^o$
Water	1.35	0.64	0.86
Ethylene glycol (20%)	1.30	0.67	0.86
$CaCl_2$ ($I = 1.6$)	1.22	0.65	0.79
$(NH_4)_2SO_4$ ($I = 1.6$)	1.22	0.64	0.78
DMSO (20%) + $CaCl_2$	1.07	0.74	0.80

$\gamma_2^o\Gamma_2/\gamma_2/\gamma_3^o\Gamma_3/\gamma_3$ is approximately constant when the product amino acid (2) is crystallized from *any* aqueous solution containing both the product (2) and impurity (3), provided the solutions are dilute.

This conclusion is confirmed by examining the purity data in terms of Eq. 8. Values of the relative distribution of *L*-isoleucine crystallized in the presence of *L*-leucine at 297 K are presented in Table 3, and for *L*-leucine crystallized in the presence of *L*-isoleucine at 297 K are presented in Table 4. Also presented are the relative pure-component solubilities of the two amino acids at 297 K, and the product of the relative distribution and the relative solubility at 297 K. The last column in Table 3 shows that the average value of $\beta_{32}\alpha_{32}^o$, and hence $\gamma_2^o\Gamma_2/\gamma_2/\gamma_3^o\Gamma_3/\gamma_3$, is 0.82 in the five systems in which *L*-isoleucine is the product and *L*-leucine is the impurity. The last column in Table 4 shows that the activity-coefficient ratio has an average value of approximately 0.42 in the two systems in which *L*-leucine is the product and *L*-isoleucine is the impurity, again irrespective of the solvent. This lends further verification to the hypothesis that the purity is directly related to the relative solubility of the pure species.

The activity-coefficient ratio calculated from one system (say, a system with pure water as the solvent) can thus be used to predict the purity of *L*-isoleucine and *L*-leucine crystals grown from other solvent mixtures if the pure amino acid solubilities in the same solvents are available. This should also be true for any two isomorphic or near-isomorphic compounds crystallized from their solutions in any solvent, provided the solutions are dilute.

Discussion

The success of the predictive model depends on the observation that the activity coefficient ratio $\gamma_2^o\Gamma_2/\gamma_2/\gamma_3^o\Gamma_3/\gamma_3$ or the product $\beta_{32}\alpha_{32}^o$ is approximately constant for a particular product–impurity pair (amino acids in the present study), irrespective of the solvent. If the dilute-solution assumption is valid for these solutions, as suggested by the solubility data, then the liquid-phase activity coefficients cancel and the solid-phase activity-coefficient behavior would have to be of the type shown in Figure 6 for the ratio Γ_3/Γ_2 to remain constant over a (small) range of concentrations. Alternatively, if there is an enhancement in solubility when two or more amino acids are present in solution, the change in the activity-coefficient ratio in solution may cancel the small change in the activity-coefficient ratio in the solid phase. Further research is required to determine whether the dominant contribution is from liquid-phase or solid-phase nonidealities. Additional research is also required to determine why $\gamma_2^o\Gamma_2/\gamma_2/\gamma_3^o\Gamma_3/\gamma_3$ is apparently constant for two amino acids in different aqueous solvents. It may be due to the fact that the solutions are dilute, and the effect of changing the solution environment on the two amino acids is similar. Certainly, the addition of cosolvents or salts leads to significant changes in solubility, but little change in the solubility ratio for two isomorphic amino acids (Table 3). What is clear, however, is that both the solid and the liquid solutions exhibit nonideal solution behavior.

Table 4. Relative Distributions and Relative Solubilities when *L*-Leucine (component 2) is Crystallized from Aqueous Systems in the Presence of *L*-Isoleucine (component 3) at 297 K

Aqueous System	β_{32}	α_{32}^o	$\beta_{32}\alpha_{32}^o$
Water	0.26	1.56	0.41
DMSO (20%) + $CaCl_2$	0.31	1.35	0.42

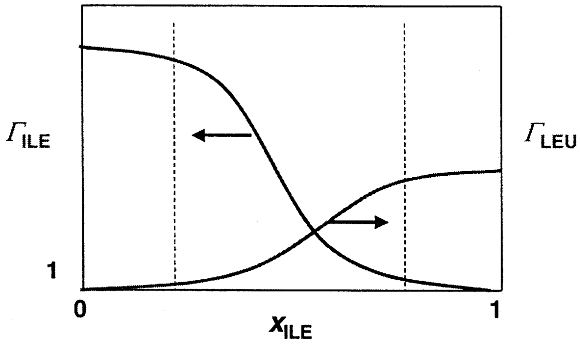

Figure 6. Postulated solid–liquid behavior of *L*-leucine + *L*-isoleucine mixtures.

Conclusions

Thermodynamic relationships were derived for solid–liquid equilibrium behavior involving either pure or mixed solids in equilibrium with their dilute solutions. The pure-solid solubility relationship was used to correlate data for *L*-leucine, *L*-isoleucine, and *L*-valine in a variety of solvents with excellent results. The results were used to correlate the purity of *L*-isoleucine crystals grown from aqueous solutions of *L*-isoleucine and *L*-leucine. With this information, and the pure-component solubilities in other aqueous solutions, it proved possible to predict purity enhancement or deterioration in crystals grown from aqueous solutions with cosolvents and/or additives. The product of the relative distribution and the relative solubility was found to be constant for a pair of product and impurity amino acids, irrespective of the solvent.

Acknowledgment

The authors thank Professor John Prausnitz for many helpful comments.

Notation

A, B = parameters in Eq. 8
A', B', C' = parameters in Eq. 9
ΔC_p = difference in heat capacity between the liquid and solid phases

f = fugacity

Δh^{fus} = enthalpy of fusion

I = ionic strength, mol kg^{-1}

R = gas constant

T = temperature

T_t = triple-point or melting temperature

x = mole fraction in the liquid phase

z = mole fraction in the solid phase

Greek letters

α = relative solubility

β = relative distribution of solutes

γ = activity coefficient in liquid phase

Γ = activity coefficient in solid phase

Superscripts and subscripts

L = liquid phase

S = solid phase

$*$ = standard state

o = pure solute property

∞ = infinite dilution

$2, 3$ = solutes

Literature Cited

Chen, C. C., Y. Zhu, and L. B. Evans, "Phase Partitioning of Biomolecules: Solubilities of Amino Acids," *Biotechnol. Prog.*, **5**, 111 (1989).

Givand, J. C., A. S. Teja, and R. W. Rousseau, "Manipulating Crystallization Variables to Enhance Crystal Purity," *J. Cryst. Growth*, **199**, 1340 (1999).

Givand, J. C., A. S. Teja, and R. W. Rousseau, " Effect of Relative Solubility on Amino Acid Crystal Purity," *AIChE J.*, **47**, 2705 (2001).

Givand, J. C., B.-K. Chang, A. S. Teja, and R. W. Rousseau, "Distribution of Isomorphic Amino Acids Between a Crystal Phase and an Aqueous Solution," *Ind. Eng. Chem. Res.*, **41**, 1873 (2002).

Gupta, R. B., and R. A. Heidemann, "Solubility Models for Amino Acids and Antibiotics," *AIChE J.*, **36**, 333 (1990).

Orella, C. J., and D. J. Kirwan, "The Solubility of Amino Acids in Mixtures of Water and Aliphatic Alcohols," *Biotechnol. Prog.*, **5**(3), 89 (1989).

Orella, C. J., and D. J. Kirwan, "Correlation of Amino Acid Solubilities in Aqueous Aliphatic Solutions," *Ind. Eng. Chem. Res.*, **30**, 1040 (1991).

Peres, A. M., and E. A. Macedo, "Representation of Solubilities of Amino Acids Using the UNIQUAC Model for Electrolytes," *Chem. Eng. Sci.*, **49**, 3803 (1994).

Pinho, S. P., C. M. Silva, and E. A. Macedo, "Solubility of Amino Acids: A Group Contribution Model Involving Phase and Chemical Equilibria," *Ind. Eng. Chem. Res.*, **33**, 1341 (1994).

Prausnitz, J. M., R. N. Lichtenthaler, and E. G. Azevedo, *Molecular Thermodynamics of Fluid-Phase Equilibria*, 3rd ed., Prentice Hall, Englewood Cliffs, NJ (1999).

Rosenberger, F., and H. G. Riveros, "Segregation in Alkali Halide Crystallization from Aqueous Solutions," *J. Chem. Phys.*, **60**, 668 (1974).

Manuscript received Jan. 28, 2002, and revision received May 21, 2002.