Temperature Dependence of Solubility Predicted from Thermodynamic Data Measured at a Single Temperature: Application to $\alpha$, $\beta$, and $\gamma$-Glycine

Andrew Manson, Jan Sefcik, and Leo Lue*  

**ABSTRACT:** Understanding of solid–liquid equilibria for polymorphic systems is crucial for rational design and efficient operation of crystallization processes. In this work, we present a framework to determine the temperature dependent solubility based on experimentally accessible thermodynamic data measured at a single temperature. Using this approach, we investigate aqueous solubility of $\alpha$, $\beta$, and $\gamma$-glycine, which, despite numerous studies, have considerable quantitative uncertainty, in particular for the most stable ($\gamma$) and the least stable ($\beta$) solid forms. We benchmark our framework on $\alpha$-glycine giving predictions in excellent agreement with direct solubility measurements between 273−340 K, using only thermodynamic data measured at the reference temperature (298.15 K). We analyze the sensitivity of solubility predictions with respect to underlying measurement uncertainty, as well as the excess Gibbs free energy models used to derive required thermodynamic quantities before providing solubility predictions for $\beta$ and $\gamma$-glycine between 273−310 and 273−330 K, respectively. Crucially, this approach to predict solubility as a function of temperature does not rely on measurement of solute melting properties which will be particularly useful for compounds that undergo thermal decomposition or polymorph transition prior to melting.

**INTRODUCTION**

Glycine is the most simple amino acid known to crystallize in one of three polymorphs identified as $\gamma$, $\alpha$, and $\beta$ (in order of decreasing thermodynamic stability at ambient conditions). Different crystallization techniques (i.e., cooling, evaporation, antisolvent, etc.) can be used to crystallize respective glycine polymorphs while additives (e.g., salts) and process conditions (e.g., stirring) can also be used to change the polymorphic outcome. Rational design and efficient operation of crystallization processes requires quantitative understanding of solid–liquid equilibria over a suitable range of operating conditions, as equilibrium solubility compositions and respective supersaturations need to be known for relevant solid forms.

The aqueous solubility of glycine polymorphs has been a topic of numerous investigations (see refs 2 and 4 for reviews). As pointed out in recent works, aqueous solubility data for $\alpha$- and $\gamma$-glycine is largely inconsistent. Hence, direct measurements of the aqueous solubility of glycine polymorphs using conventional techniques appear challenging for the most stable solid form, $\gamma$-glycine, and even for $\alpha$-glycine at elevated temperatures (above 330 K). This may be related to issues with mass transfer and particularly with very slow growth (and perhaps dissolution) rates at small supersaturations in highly concentrated solutions, so that it could be challenging to ensure that solid–liquid equilibrium is reached. For $\beta$-glycine, there is currently limited solubility data available, most likely due to rapid recrystallization to $\alpha$-glycine. In turn, this then motivates our work investigating alternative approaches to prediction of solubilities using other thermodynamic data that may be accessible.

Measurements of thermodynamic properties other than solubility (e.g., enthalpy of solution, vapor pressure, etc.) exist over a fairly broad range of temperatures for the glycine–water system; however, the vast majority of measurements are reported around 298.15 K between dilute to moderate concentrations of glycine. Practical difficulties in working with glycine–water solutions at higher temperatures and moderate concentrations severely limit the availability of experimental data at elevated temperatures.

Thermodynamic models provide a means by which experimental measurements (e.g., solubility, heat capacity, etc.) can be used to make estimates beyond the specific conditions at which the measurements were performed and for properties other than those that were measured. In this approach, mathematical expressions are used to represent the free energy (e.g., Gibbs or Helmholtz free energy) of the

Received: October 18, 2021  
Revised: January 25, 2022  
Published: February 8, 2022
system, and from this all thermodynamic properties of the system can be determined. Thermodynamic models are attractive, in theory, because they can greatly reduce the experimental effort required, for example, to determine the value of a specific physical property of a system over a broad range of conditions.

The theoretical basis and complexity of thermodynamic models can vary significantly, in part, depending on their application. At one extreme, there are fully empirical models with parameters that are adjusted to reproduce experimental data but which do not have any direct physical interpretation. For example, when solubility is assumed to be a polynomial function of temperature and the coefficients are regressed against a series of direct solubility measurements at different temperatures. The resulting model is used to predict solubility at conditions not covered by experiment through interpolation and extrapolation. While this approach can be quite useful in practical applications, its reliability outside the range of conditions covered by the experimental data is uncertain.

At the other extreme is the use of free energy models that are firmly rooted in molecular statistical mechanics with physical meaningful parameters that are related to properties such as molecular sizes or interactions strengths. This allows the a priori estimate of the values of the parameters, or at least provide physical bounds as to their possible range, without need for experimental input. While these more theoretically based approaches are expected to have a greater range over which the results can be reliably extrapolated, there is almost always a need to adjust the model parameters, due to approximations and simplifications made during their derivation, in order to quantitatively reproduce experimentally measured properties.

Within this approach, for the prediction of solubility, a thermodynamic model is required for the liquid phase, and a separate model is required for the solid phase. For glycine–water mixtures, the Pitzer,\textsuperscript{10} modified mean spherical approximation (MSA),\textsuperscript{10} and PC-SAFT\textsuperscript{12,13} as well as many other, models have been used successfully to describe the liquid phase. Typically, the parameters of these models need to be regressed across a large collection of thermodynamic measurements over a wide range of conditions. The predictions of these models can then depend sensitively on the precise forms chosen for these parameters, such as how they depend on temperature and composition. What is needed is a manner to assess how far an experimental measurement can be reliably extrapolated.

In addition, a thermodynamic model is required for the solid, which typically rely on pure solute melting enthalpy, melting temperature and heat capacity (solid and liquid solution reference state, typically taken as the pure liquid solute).\textsuperscript{8,14} However, this approach is limited for systems where melting properties are not experimentally accessible using conventional approaches, due to thermal decomposition (as is the case for glycine), polymorph transition, etc. However, it should be noted that recently, novel approaches have been reported recently that mitigate these issues.\textsuperscript{13} Regardless, simplifications employed during model derivation can lead to significant errors in solubility predictions at temperatures far from the solute melting temperature.

In this work, we develop a novel approach to determine the temperature dependence of solubility, which, in principle, relies entirely on thermodynamic data collected at a single temperature by approximation of parameter dependencies through first-order Taylor series expansions. This is advantageous in the sense that it limits the amount of thermodynamic data required to make solubility predictions. With this approach, we provide an estimate for the aqueous solubility of α-glycine, from its eutectic point to the solid-solution–vapor triple point (i.e., intersection of solubility and boiling curve). This is based on enforcing consistency with thermodynamic measurements for a broad range of properties at 298.15 K. For β and γ-glycine, where solubility measurements are inconsistent or limited, we develop methods for estimating the solubility at 298.15 K, as well as its dependence on temperature, using ancillary thermodynamic data.

The remainder of the Article is structured as follows: First, the relevant thermodynamic theory is introduced. Then, the relevant approximation is described and an interpretation of model parameters is provided in the context of experimental measurements. Then, a review of thermodynamic data (including solubility, activity, enthalpy, and heat capacity) and results from excess free energy model fitting are given, followed by solubility predictions for α-glycine. The sensitivity of our approach to uncertainty in experimental data and excess free energy model choice is then discussed. Finally, the extension of the method to β- and γ-glycine is presented.

**THEORY**

We consider a solute (denoted as 1), which is dissolved in a solvent (denoted as 2). At sufficiently high concentrations, the solute precipitates as a pure solid. The temperature dependence of the mole fraction of the solute at saturation can be expressed by the following relation:

\[
\frac{d \ln x_{1,\text{sat}}(T)}{dT} = \frac{1}{RT^2} \left[ \frac{\tilde{R}_1(T, x_{1,\text{sat}}(T)) - h_1^\circ(T)}{1 - (1 - x_{1,\text{sat}}(T))\partial \ln \gamma_2(T, x_{1,\text{sat}}(T))/\partial x_1} \right]
\]

(1)

where \( R \) is the gas constant, \( \tilde{R}_1(T, x_{1,\text{sat}}(T)) \) is the partial molar enthalpy of the solute in a saturated solution at temperature \( T \) and \( h_1^\circ(T) \) is the molar enthalpy of the pure solid at \( T \) (which is not a function of solution composition). The solute activity coefficient \( \gamma_2 \) is defined with the chemical potential of the solute written as \( \beta \mu_1^\circ = \beta \mu_1^\circ(T) + \ln x_2 \gamma_2 \), where \( \mu_1^\circ(T) \) is the chemical potential of pure liquid solute at temperature \( T \). The quantity \( \tilde{R}_1(T, x_{1,\text{sat}}(T)) - h_1^\circ(T) \) is sometimes referred to as the “differential heat of solution”, and the quantity \((1 - x_{1,\text{sat}}(T))\partial \ln \gamma_2(T, x_{1,\text{sat}}(T))/\partial x_1 \) is related to the variation of the solvent activity coefficient (defined through the relation \( \beta \mu_2^\circ = \beta \mu_2^\circ(T) + \ln x_2 \gamma_2 \), where \( \mu_2^\circ(T) \) is the chemical potential of pure solvent at temperature \( T \) with composition, evaluated at saturation temperature and composition. For convenience, we refer to \( \tilde{R}_1(T, x_{1,\text{sat}}(T)) - h_1^\circ(T) \) as \( a \), and \((1 - x_{1,\text{sat}}(T))\partial \ln \gamma_2(T, x_{1,\text{sat}}(T))/\partial x_1 \) as \( b \). In principle, both \( a \) and \( b \) are experimentally accessible, and if they were known in addition to a single solubility point, the solubility at any temperature could be predicted.

Expressions equivalent to eq 1 have been reported by various authors,\textsuperscript{13,15-18} however, it can be derived by considering that, at thermodynamic equilibrium, the chemical potential of a
pure solid solute is equal to the chemical potential of the solute in a saturated solution (at the same temperature):

\[ \beta \mu^i_1(T) = \beta \mu^i(T, x_{1,\text{sat}}(T)) \]  

(3)

where \( \beta = (RT)^{-1} \), \( R \) is the gas constant, \( \mu^i_1(T) \) is the chemical potential of pure solid solute and \( \mu^i(T, x_{1,\text{sat}}(T)) \) is the chemical potential of the solute in a saturated solution. Equation 3 is valid at any point along the solubility curve. More generally, small changes in the chemical potential terms can be expressed as

\[ d\beta \mu^i_1(T) = \frac{\partial \mu^i_1(T)}{\partial \beta} d\beta \]  

(4)

\[ d\beta \mu^i_1(T, x_i) = \frac{\partial \mu^i_1(T, x_i)}{\partial \beta} d\beta + \frac{\partial \mu^i_1(T, x_i)}{\partial \ln x_i} d\ln x_i \]  

(5)

which results from the fact that they are continuous functions of \( T \) and \( x_i \). Combining eqs 4 and 5, while noting that eq 3 can be expressed as \( d\beta \mu^i_1(T, x_{1,\text{sat}}(T)) = d\beta \mu^i(T, x_{1,\text{sat}}(T)) \), gives

\[ \frac{\partial \mu^i_1(T, x_{1,\text{sat}}(T))}{\partial \ln x_i} = \frac{d\ln \chi(T, x_{1,\text{sat}}(T))}{\partial \ln x_i} \]  

(6)

which, again, is valid along the solubility curve. Going further

\[ \frac{d\ln x_{1,\text{sat}}(T)}{dT} = \frac{\frac{\partial}{\partial x_{1,\text{sat}}}(\chi(T, x_{1,\text{sat}}(T)) - h^i_1(T))}{1 + \frac{\partial}{\partial x_{1,\text{sat}}}(\chi(T, x_{1,\text{sat}}(T)))/\partial \ln x_i} \]  

(7)

where we make use of the Gibbs–Helmholtz relation \( h^i_1 = \beta \mu^i / \beta \), where \( h^i_1 \) is the partial molar enthalpy of species \( i \). Focusing on \( \partial \mu^i_1(T, x_{1,\text{sat}}(T))/\partial \ln x_i \) gives

\[ \frac{\partial \mu^i_1(T, x_{1,\text{sat}}(T))}{\partial \ln x_i} = \left[ 1 + \frac{\partial \ln \chi(T, x_{1,\text{sat}}(T))}{\partial \ln x_i} \right] \]  

(8)

which, coupled with eq 7,

\[ \frac{d\ln x_{1,\text{sat}}(T)}{dT} = \frac{\frac{\partial}{\partial x_{1,\text{sat}}}(\chi(T, x_{1,\text{sat}}(T)) - h^i_1(T))}{1 + \frac{\partial}{\partial x_{1,\text{sat}}}(\chi(T, x_{1,\text{sat}}(T)))/\partial \ln x_i} \]  

(9)

which is an expression identical with that reported previously.\(^{15-18}\) Since data relating to the activity of solvent in solution is more widely reported, we use the Gibbs–Duhamel relationship to replace the solute activity coefficient term:

\[ \frac{\partial \ln \chi}{\partial \ln x} = \frac{x_1}{x_{1,\text{sat}}} \frac{\partial \ln \chi}{\partial \ln x} = -x_1 \frac{\partial \ln \chi}{\partial \ln x} + x_1 \frac{\partial \ln \chi}{\partial \ln x} \]  

noting that we make the arbitrary choice to use the partial derivative with respect to \( x_{1,i} \) rather than \( \ln x_{1,i} \), which when combined with eq 9 gives eq 1. For interested readers, we report another derivation in our Supporting Information, based entirely on earlier work found in ref 17. As an aside, eq 1 is sometimes reported in the form

\[ \frac{d\ln x_i}{dT} = \frac{\Delta H^{\text{vH}}}{RT^2} \]  

(11)

where \( \Delta H^{\text{vH}} \) is referred to as the “van’t Hoff enthalpy of solution.”\(^{18,19}\)

The utility of eq 1 is 2-fold. First, if the solubility of a compound is known as a function of temperature, the right side can be determined, which directly leads to an estimate of the differential heat of solution and related solution properties.\(^{17,20}\) Alternatively, if the right side were known as a function of temperature and solution composition, the solubility as a function of temperature could be calculated by integration of the equation, given the solubility at one temperature as a starting point.\(^{21}\)

In general, \( a \) and \( b \) in eq 2 are functions of both temperature and composition; however, it is unclear what those dependencies are. Ideally, we could collate sufficient experimental data to develop functional expressions for both parameters (i.e., across the entire range of temperature and solubility); however, in practice this would require significant effort. In addition, it is unclear if techniques are available that allow evaluation of the required measurements at extremes of temperature and solution composition.

In the absence of sufficient data to correlate functions for model parameters, we can approximate their temperature and composition dependence using Taylor’s theorem, which is a multivariate technique used to locally approximate analytic, multivariate functions in terms of their partial derivatives evaluated at a chosen reference point.

For convenience, we apply the first-order approximation of Taylor’s theorem, which, for a two-variable continuous function, is given as

\[ f(x, y) \approx f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0) \]  

(12)

\[ \approx f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0) \]  

(13)

It should be noted that the first-order expansion will limit the accuracy of approximation far from the chosen reference point. For the purpose of this work, we perform the following expansions for \( a \) and \( b \), respectively:

\[ a(T, x_i) \approx a_0(T_0, x_{1,0}) + a_1 \left( \frac{T}{T_0} - 1 \right) + a_2 \left( \frac{x_1}{x_{1,0}} - 1 \right) \]  

(14)

\[ b(T, x_i) \approx b_0(T_0, x_{1,0}) + b_1 \left( \frac{T}{T_0} - 1 \right) + b_2 \left( \frac{x_1}{x_{1,0}} - 1 \right) \]  

(15)

where all coefficients are defined in the Supporting Information.

We have chosen to approximate \( a(T, x_i) \) in terms of \( T \) and \( x_i \), since \( d\beta /dT = -\epsilon \) which makes it possible to use heat capacity data to evaluate \( a_0 \), while we have chosen to approximate \( b(\beta, x) \) in terms of \( \beta \) and \( x \) because it allows us to relate \( a_2 \) and \( b_1 \). However, it should also be noted that the expansion can be performed in many other ways.
Substituting eqs 14 and 15 into eq 2, the solubility relation becomes
\[
\frac{d \ln x_{\text{sat}}(T)}{dT} 
\approx \frac{1}{RT^2} a(T) + a_0(T) - 1 + a_1 \left( \frac{x_{\text{sat}}(T)}{x_{\text{sat}}(T_0)} - 1 \right)
\]
from which the solubility at any temperature can be predicted by numerical integration if all model coefficients and the solubility at \( T_0 \) is known.

For notational convenience, the composition dependence of coefficients in eq 16 will be represented by \( x_1 \) and \( x_{1,0} \) in place of \( x_{\text{sat}}(T) \) and \( x_{\text{sat}}(T_0) \), respectively, throughout the remainder of the paper. However, it should be emphasized that these refer to the quantities evaluated at or along the solubility curve.

**Relationship Between Model Coefficients and Measured Data.** It is possible to relate each of the coefficients in eq 1 to various experimentally measurable thermodynamic quantities. Depending on the coefficient, the required thermodynamic quantities may be accessible through a direct measurement (e.g., enthalpy of solution to infinite dilution) or by regressing an excess Gibbs free energy model to different types of measurements and using the model to derive necessary quantities (e.g., using vapor pressure data to evaluate \( \partial \ln \gamma(x_1) / \partial x_1 \)). It should be noted that there are many ways in which experimental data could be used to evaluate the coefficients in eq 16, resulting from the many relationships between thermodynamic quantities. However, for the purpose of this work, we relate coefficients to experimental measurements performed at a single temperature. For convenience, all \( a \) coefficients and all \( b \) coefficients are discussed together.

**Determination of \( a \) Coefficients.** The \( a \) coefficient can be expanded as
\[
a(T, x_{\text{sat}}(T)) \approx \Delta h_1^{\infty, x_1}(T_0) + T_0 \left( \frac{T}{T_0} - 1 \right) + x_{\text{sat}}(T_0) \left( \frac{\partial \Delta h_1^{(0)}}{\partial x_1}(T_0) \frac{x_{\text{sat}}(T)}{x_{\text{sat}}(T_0)} - 1 \right)
\]
where \( \Delta h_1^{(0)} = \Delta h_1(T_0, x_{\text{sat}}(T_0)) - h_1^{\infty}(T_0) \) could be evaluated directly if “differential heat of solution” measurements were available; however, in the absence of these measurements, it can be shown that
\[
\Delta h_1^{(0)} = \Delta h^{\infty, x_1}(T_0) + \frac{\partial \ln \gamma_1^{x_1}(T_0)}{\partial \beta}
\]
where \( \Delta h^{\infty, x_1} \) is the enthalpy of solution when forming an “infinitely dilute” solution. It should be noted that we introduce \( \gamma_1^{x_1} = \gamma_1 / \gamma_1^{x_1} \) as the solute activity coefficient defined relative to an infinitely dilute solution, where \( \gamma_1^{x_1} = \lim_{x_1 \to 0} \gamma_1 \). The quantity \( \partial \ln \gamma_1^{x_1} / \partial \beta = \partial \ln \gamma_1(T_0, x_{\text{sat}}(T_0)) / \partial \beta \) can be evaluated by regressing an excess Gibbs free energy model to “enthalpy of dilution” data (for cases where the enthalpy of dilution is given per mole of solute in solution) via
\[
\frac{\Delta H_{\text{dd}}(T, x^f, x^c)}{n_1} = \left( \sum_{i} n_i \frac{\partial \ln \gamma_i(T, x)}{\partial \beta} \right)_{x^i = x^f} - \left( \sum_{i} n_i \frac{\partial \ln \gamma_i(T, x)}{\partial \beta} \right)_{x^i = x^c}
\]
In addition, since \( \Delta h^{\infty, x_1}(T_0) \) is independent of composition, it can be shown that
\[
\frac{\partial \Delta h_1^{(0)}}{\partial x_1} = \frac{\partial^2 \ln \gamma_1^{x_1}(T_0)}{\partial x_1^2} \partial \beta
\]
which can be evaluated from the same excess Gibbs free energy model regressed to enthalpy of dilution data. Given that \( c_p = \partial h / \partial T \), it can be shown that
\[
\frac{\partial \Delta h_1^{(0)}}{\partial T} = \Delta c_p^{(0)}
\]
where \( \Delta c_p^{(0)} = \frac{c_p}{\gamma_1^{x_1}(T_0, x_{\text{sat}}(T_0))} - c_p^{x_1}(T_0) \), \( c_p^{x_1} \) is the molar heat capacity of crystalline glycine, which can be measured directly, and \( c_p^{x_1} \) is the partial molar heat capacity of glycine in solution which can be interpreted as
\[
c_p^{x_1}(T_0, x_{\text{sat}}(T_0)) = c_p^{x_1}(T_0) + \frac{\partial^2 \ln \gamma_1^{x_1}(T_0)}{\partial T \partial \beta}
\]
which, given an excess Gibbs free energy model, can be estimated by regressing solution heat capacity data, since
\[
c_p^{\text{soln}}(T, x) = \sum_{i} x_i c_p^{x_1}(T) + \sum_{i} x_i \frac{\partial^2 \ln \gamma_i(T, x)}{\partial T \partial \beta}
\]
**Determination of \( b \) Coefficients.** In the context of this work:
\[
b(\beta, x_{\text{sat}}(T)) \approx x_{2, \text{sat}}(T_0) \left( \frac{\partial \ln \gamma_2}{\partial x_1} \right)_{x_1 = x_{\text{sat}}(T_0)} + \beta_0 \left( \frac{\partial \ln \gamma_2}{\partial x_1} \right)_{x_1 = x_{\text{sat}}(T_0)} \left( \frac{T}{T_0} - 1 \right) + x_{\text{sat}}(T_0) \left( \frac{\partial x_{2, \text{sat}}(T_0)}{\partial x_1} \right) \left( \frac{\partial \ln \gamma_2}{\partial x_1} \right)_{x_1 = x_{\text{sat}}(T_0)} \left( \frac{x_{\text{sat}}(T)}{x_{\text{sat}}(T_0)} - 1 \right)
\]
where again \( \ln \gamma_2^{x_1} = \ln \gamma_2(T_0, x_{\text{sat}}(T_0)) \) is introduced for notational convenience. It can be shown that
\[
\frac{\partial \Delta h_1^{(0)}}{\partial x_1} \left( \frac{\partial \ln \gamma_2^{x_1}(T_0, x_{\text{sat}}(T_0))}{\partial x_1} \right) = -\frac{\partial \ln \gamma_2^{x_1}(T_0, x_{\text{sat}}(T_0))}{\partial x_1} + x_{2, \text{sat}}(T_0) \left( \frac{\partial^2 \ln \gamma_2^{x_1}(T_0, x_{\text{sat}}(T_0))}{\partial x_1^2} \right)_{x_1 = x_{\text{sat}}(T_0)}
\]
which can be evaluated by regressing an excess Gibbs free energy model to “activity” related measurements (i.e., vapor pressure, isopiestic molality, etc.) taken at the reference temperature and at various compositions, and
\[
\frac{\partial}{\partial \beta} \left( x_{2, \text{sat}}(T_0) \frac{\partial \ln \gamma_2^{x_1}}{\partial x_1} \right) = -x_{1, \text{sat}}(T_0) \left( \frac{\partial^2 \ln \gamma_2^{x_1}}{\partial x_1^2} \right)_{x_1 = x_{\text{sat}}(T_0)}
\]
Excess Gibbs Free Energy Models. Several parameters required to estimate the coefficients in eq 16 are derivatives of continuous thermodynamic quantities. Given the discrete nature of many experiments, estimation of required parameters requires correlation of data using models. In the context of thermodynamics, the excess Gibbs free energy is a concept that each of the required quantities can be related to. Historically, many attempts have been made to derive physically interpretable excess Gibbs free energy models and as such, for the purpose of this work, we use the Scatchard–Hildebrand and Scatchard–Hildebrand–Flory–Huggins excess Gibbs free energy models. 23

Scatchard–Hildebrand. The Scatchard–Hildebrand excess free energy model (for a binary solution comprising glycine and water) is given by

$$\beta G^e = (n_1v^L_1 + n_2v^L_2)\phi_i\phi_j\chi'$$  \hspace{1cm} (27)

where $v^L_i$ is the molar volume of component $i$, $\phi_i$ is the volume fraction of component $i$, and $\chi'$ is a binary interaction parameter. This gives

$$\ln \gamma_2 = \beta v^L_1\phi_1^2\chi'$$  \hspace{1cm} (28)

$$\ln \gamma_1 = \beta v^L_1\phi_2^2\chi'$$  \hspace{1cm} (29)

For convenience, we define $\chi(\beta) = \chi'\beta v^L_1$ and $v^* = v^L_1/v^L_2$, which gives

$$\ln \gamma_2 = \phi_1^2\chi$$  \hspace{1cm} (30)

$$\ln \gamma_1 = v^*\phi_2^2\chi$$  \hspace{1cm} (31)

where we have assumed $\chi$ to be a function of temperature, while the molar volumes (i.e., $v^*$) are independent of temperature and composition. These assumptions make it straightforward to use the same model throughout the correlation of different thermodynamic data. In addition

$$\phi_i = \frac{x_i v^*}{x_i v^* + x_2}$$  \hspace{1cm} (32)

![Figure 1. "Van’t Hoff" style plot of literature solubility data for $\alpha$, $\beta$, $\gamma$, and undefined glycine between 270–430 K. Note that $T_0 = 298.15$ K. The black dashed line is the $\alpha$-glycine solubility estimate from ref 2 and is shown in all plots as a guide.](https://doi.org/10.1021/acs.cgd.1c01217)
The excess enthalpy of this model, assuming $\chi'$ is a function of temperature only and $v^R$ is independent of temperature and composition, is given by

$$H^E = n_1v^L_1\phi_1\phi_1'x_1' + n_2v^L_2\phi_2\phi_2'x_2'$$

and the excess heat capacity is given by

$$C_p^E = n_1v^L_1\phi_1\phi_1'x_1' + n_2v^L_2\phi_2\phi_2'x_2'$$

where $x_1' = \frac{\partial x_1}{\partial T}\phi_1$, and $x_2' = \frac{\partial x_2}{\partial T}\phi_2$.

Scatchard–Hildebrand–Flory–Huggins. To account for differences in molecular size we can add a “Flory–Huggins” term to the Scatchard–Hildebrand free energy expression; which, for a binary solution, is given by

$$\beta G^E = (n_1v^L_1 + n_2v^L_2)\phi_1\phi_1' + n_1\ln\frac{\phi_1}{x_1} + n_2\ln\frac{\phi_2}{x_2}$$

which gives

$$\ln x_1 = \phi_1^2 x_1 + \left(\frac{\phi_2}{x_2} + 1 - \frac{\phi_2}{x_1}\right)$$

$$\ln x_2 = n^R\phi_2^2 x_2 + \left(\frac{\phi_1}{x_1} + 1 - \frac{\phi_1}{x_2}\right)$$

It should be noted that $H^E$ and $C_p^E$ are the same for both the Scatchard–Hildebrand–Flory–Huggins and Scatchard–Hildebrand when $x'$ is a function of temperature only and $v^R$ is a constant independent of temperature and composition, which is an assumption used in this work.

## RESULTS AND DISCUSSION

### Solubility Review

There are numerous reports of glycine–water solubility measurements in the literature. Some articles present data for an identified polymorph (see ref 24 for $\alpha$, $\beta$, and $\gamma$, refs 25–29 for $\alpha$ and $\gamma$, ref 32 for $\alpha$, and ref 33–39), which is problematic, since crystal polymorphs have different solubility in the same solvent at the same temperature and pressure.

Some of this data has been reviewed recently, however, both reviews have specific areas of focus and limitations. Datta et al. make no distinction between glycine polymorphs, choosing to review only data that has been identified as “glycine”. Their review does not include any data labeled with glycine polymorphs, excluding a number of data sets from their review.

Rowland reviewed solubility data as part of their work on an equation of state model parameters, as well as the standard state properties for the glycine–water solution. Rowland’s review is restricted to data assumed to be $\alpha$-glycine, which includes data points reported as $\alpha$-glycine, as well as data where the glycine polymorph is unspecified; however, $\beta$ and $\gamma$-glycine were out of the scope of the work.

The primary aim of our glycine–water solubility review is to consider other glycine polymorphs alongside $\alpha$-glycine, as well as any missing data not included previously, especially at elevated temperatures where glycine solubility data are sparse. We proceeded as follows: first, if the solubility data had the polymorph labeled (e.g., $\alpha$, $\beta$, or $\gamma$) it was included. Then, if the solubility data was for an undefined glycine polymorph, it was included only if the data set included data above 330 K. Data meeting either of these criteria are presented in Figure 1.

We note that some solubility compilations (e.g., see Table A5 in ref 7) show glycine solubility up to 373 K, without providing a source reference. However, we believe data in ref is based on an extrapolation of lower temperature data as reported early on, this was also before polymorphism of glycine was established.

From Figure 1, each data set shows the solubility of glycine in water increasing with temperature as generally expected. For data labeled as either $\alpha$ or $\gamma$-glycine, there are a similar number of data sets because of various studies reporting measurements for both polymorphs together. In addition, it can be seen that data is sparse above 350 K, with only one data set for $\alpha$ and $\gamma$-glycine, respectively. Interestingly, the labeled polymorph data sets have greater variability, while the undefined polymorph data sets appear more consistent. It is unclear why this is the case; however, following our own assessment and in agreement with Rowland, we attribute the undefined data sets as $\alpha$-glycine. On the basis of the slope of their data, we observe that data from Devi and Igarashi for both $\alpha$ and $\gamma$-glycine is inconsistent with each other and other data reported in the literature.

Based on our review, there is only one report of $\beta$-glycine in the literature; however, it is interesting to note reports that direct measurement of $\beta$-glycine in pure water are challenging due to kinetic instability in water and resulting rapid transformation to $\alpha$-glycine. Taking Rowland’s estimate for $\alpha$-glycine as a guide, it appears data reported as $\beta$-glycine could be $\alpha$-glycine.

### Thermodynamic Properties

As discussed above, Rowland published a comprehensive review of thermodynamic data for glycine–water mixtures. Much of the data presented there corresponds to data required to evaluate the coefficients in eq 16 and as such, the work was used as a reference for thermodynamic data. An initial review indicated that most of the thermodynamic data for glycine–water solutions has been collected at 298.15 K, as such we chose $T_0 = 298.15$ K as the start point for our Taylor expansion. Following a further review of literature, any additional data required to evaluate the coefficients in eq 16 was found elsewhere.

Solubility at the chosen reference temperature is a crucial thermodynamic property required to make predictions with the approach presented in the theory section. For $\alpha$-glycine, we have chosen to take the value reported by Rowland (i.e., 3.324 mol kg$^{-1}$ of water or 0.056 (glycine mole fraction) at 298.15 K), derived from a thermodynamically consistent, semi-empirical fit to a broad range of experimental data (including solubility), which we believe to be the best estimate available in literature.

It should be noted that, for the purpose of this work, we model solution properties separately (i.e., $\ln\gamma$ is modeled independently of $\Delta H^R/n$ and $C_p^R$), rather than developing a model to simultaneously describe all thermodynamic data. Development of a comprehensive model is out with the scope of this work, and the approach described here is computationally convenient and ensured experimental data was modeled accurately.

In the following section, thermodynamic data used to evaluate the coefficients in eq 16 is described. This includes measurements from which the solvent activity coefficient can
be derived, enthalpy of dilution, and solution heat capacity, as well as measurements for enthalpy of solution to infinite dilution and crystalline glycine polymorph heat capacity (both at 298.15 K).

**Activity Coefficient Data, Fits, and Derived Properties.** From the review published by Rowland, 12 data sets from Refs. 10,12,41−50 were found from which the water activity coefficient as a function of solution composition can be derived. For seven of the data sets, the water activity coefficient \( \gamma_w \) was derived from isopiestic or osmotic vapor pressure measurements and subsequent estimation of the osmotic coefficient in glycine−water solutions of various compositions at 298.15 K using

\[
\Phi = -\frac{1}{M_w m_g} \ln(\gamma_w x_w)
\]

(39)

where \( x_w \) is the mole fraction of water, \( M_w = 0.018 \text{ kg mol}^{-1} \) is the molecular weight of water, and \( m_g \) is glycine molality. Four data sets reported values for water activity \( a_w \) at various glycine−water solution concentrations (at 298.15 K), which were converted to activity coefficients via

\[
a_w = \gamma_w x_w
\]

(40)

The final data set reported static vapor pressure measurements for glycine−water solutions of varying composition at 298.15 K. The water activity coefficient was determined from the modified Raoult’s law (assuming the vapor was ideal and comprised only of water)

\[
P = x_w \gamma_w P_w^{sat}(T)
\]

(41)

where \( P \) is the pressure and \( P_w^{sat}(T) \) is the saturated vapor pressure of pure water at 298.15 K, taken as the value provided in the data set. All data, converted to \( \ln \gamma_w \) and plotted as a function of solution composition, are given in Figure 2a.

From Figure 2a, the general trend indicates that \( \ln \gamma_w \) increases with glycine mole fraction, meaning that \( \gamma_w > 1 \). Excluding the data from Ninni, the data appears to follow the same trend; however, there is significant variability between...
datasets, particularly above \( x_g = 0.03 \). In general, evaluations made from isopiestic vapor pressure measurements appear to be more consistent in terms of internal scatter. A review of data included in Rowland’s global fit indicates that all data used was measured using an isopiestic vapor pressure technique, except from the single point that was measured using vapor pressure osmometry. For the purpose of this work, we assume the data chosen by Rowland to be the most reliable and use this to regress the Scatchard–Hildebrand and Scatchard–Hildebrand–Flory–Huggins models, respectively.

From Figure 2b, both the Scatchard–Hildebrand and Scatchard–Hildebrand–Flory–Huggins models appear to fit the activity coefficient data well. However, the choice of model has a significant impact on behavior of the first and second partial derivatives and their values at glycine polymorph solubility. For example, the first partial derivative evaluated at the estimated \( \alpha \)-glycine solubility is 0.098 and 0.075 for the Scatchard–Hildebrand and Scatchard–Hildebrand–Flory–Huggins model, respectively, while the second partial derivative is \(-0.74\) and \(-2.82\). Given that both partial derivatives are parameters used to evaluate solubility relation coefficients, the differences suggest that, in the case of discrete data, the choice of model can impact the predictive ability.

**Enthalpy of Dilution Data, Fits and Derived Properties.** Again, following Rowland there are at least 8 data sets \(^{51-58}\) available in the literature reporting enthalpy of dilution measurements. For the purpose of review, each data set was correlated by eq 19 to allow direct comparison, noting that raw measurements cannot be compared graphically. From the best fit to each data set, points corresponding to \( \partial \ln \gamma^*_g / \partial \beta \) and \( \partial^2 \ln \gamma^*_g / \partial x_g \partial \beta \) derived from Scatchard–Hildebrand best fit to data in panel a. Open symbols are values derived from fits to literature data, \(^{51-58}\) and gray dashed vertical lines indicate \( \alpha \)-glycine solubility at 298.15 K.

![Figure 3](https://cryst-growth.des.pubs.acs.org/crystal/2022/0001217/crystal.2022.0001217.html)

**Figure 3.** (a) Literature glycine–water solution enthalpy of dilution processed data at 298.15 K and (b, c) \( \partial \ln \gamma^*_g / \partial \beta \) and \( \partial^2 \ln \gamma^*_g / \partial x_g \partial \beta \) derived from Scatchard–Hildebrand best fit to data in panel a. Open symbols are values derived from fits to literature data, \(^{51-58}\) and gray dashed vertical lines indicate \( \alpha \)-glycine solubility at 298.15 K.
of this curve at the estimated solubility is required to evaluate model parameters, and thus, extrapolation is necessary, which may introduce uncertainty.

Again, eq 19 was regressed against the aggregated data set (i.e., the data sets used by Rowland), and the results are presented in Figure 3b, with \( \frac{\partial^2 \ln \gamma_g^{m}}{\partial x_g \partial \beta} \) in Figure 3c. It should be noted that our best fit estimate for \( \frac{\partial \ln \gamma_g^{m}}{\partial \beta} \) is in excellent agreement with that presented in ref 52.

**Enthalpy of Solution and Pure Crystal Heat Capacity.**

Direct measurements of the enthalpy of solution (to infinite dilution) and pure solid molar heat capacity at 298.15 K, are required to evaluate the coefficients in eq 16. Both have been reported in the literature and are presented in Table 1. We note that both \( \Delta h^{\infty,\alpha} \) and \( c_{p,298.15} \) have been measured for \( \alpha, \beta, \) and \( \gamma \)-glycine, respectively. However, we also note that each property has been measured only once and not validated by additional measurements.

**\( \alpha \)-Glycine Solubility Predictions.**

The solubility of \( \alpha \)-glycine was predicted as a function of temperature by numerically integrating eq 16 with the coefficient values...
presented in the Supporting Information, and predictions are shown in Figure 5. It should be noted that the solubility was also estimated with $a_0$ and $b_0$ only, which is referred to as the “0th-order” prediction, while predictions based on eq 16 are referred to as “1st-order” predictions.

From Figure 5, the zeroth and first-order approaches both predict an increasing solubility with temperature. From 280–330 K, irrespective of the excess Gibbs free energy model used to estimate model parameters, both zeroth and first-order approaches show good agreement in terms of solubility; however, beyond this, predictions diverge. Although solubility data is limited above 340 K, the first-order corrections shift solubility predictions toward available measurements, indicating that the first-order corrections improve those made by the zeroth-order model.

We note that our solubility predictions are consistent with the majority of available direct measurements, as shown in Figure 5. However, our predictions vary considerably compared to refs 24 and 27, shown as red open triangles. This supports initial observations detailed in our solubility review, where the slope of both data sets was inconsistent with other literature data when plotted on a van’t Hoff plot. As such, both data sets will be omitted from plots in the remainder of this work.

The effect of the excess Gibbs free energy model used to evaluate the coefficients in eq 16 is shown in Figure 5. For the zeroth-order predictions, model impact is limited—becoming significant around 380 K. However, the impact of model selection is greater for the first-order predictions. The corrected models begin to diverge at approximately 340 K. At 400 K, the solubility is predicted as 0.20 and 0.18 (glycine mole fraction) for the Scatchard–Hildebrand and Scatchard–Hildebrand–Flory–Huggins models, respectively.

Differences between the chosen excess Gibbs free energy models used to derive thermodynamic quantities are specific to quantities derived from fits to solvent activity coefficient data. Previously, it was shown that, although both models (Scatchard–Hildebrand and Scatchard–Hildebrand–Flory–Huggins) appeared to fit water activity coefficient data well, the first and second partial derivatives (with respect to glycine mole fraction) were significantly different. It is interesting to note that, despite this, both models produce consistent predictions near 298.15 K.

On the basis of the available solution thermodynamic data and solubility, we believe further progress could be made on assessing the performance of the approach detailed above in two ways. First, with more accurate high temperature solubility data, since most solubility data above 340 K comes from the same source in which solubility was determined using a novel application of “Differential Scanning Calorimetry (DSC),”

resulting in an unusual shape for the α-glycine solubility curve. Second, having further glycine–water solution thermodynamic data (e.g., water activity, enthalpy of dilution and solution heat capacity), for solution compositions in excess of 0.056 mole fraction, would be useful to ensure construction of accurate excess Gibbs free energy models and potentially allow decoupling of uncertainty introduced from the choice of excess Gibbs free energy model and eq 16, respectively.

Sensitivity Analysis. As shown above, the choice of model used to estimate coefficients in eq 16 impacted solubility predictions at temperatures far from the reference temperature; indicating an underlying model sensitivity. Given that there is limited solubility data in this region (i.e., above 340 K) it is not appropriate to assess reliability of the approach based solely on best fit parameter estimates, as these would be subject to measurement uncertainty. To account for this we perform a sensitivity analysis.

Each aggregated data set (i.e., $\ln \gamma_p, \Delta H^{\text{dil}}, n_r$, and $\gamma_{p,\text{soln}}$) used to regress model parameters was used to generate “simulated” data sets. A simulated data set is defined as a data set derived from the original aggregated data set; however, each data point was “blurred” by a random percentage, calculated from $y_{\text{sim}} = y_{\text{orig}} + y_{\text{orig}} \epsilon$, where $y$ denotes a thermodynamic quantity and $\epsilon$ is a random value drawn from a Gaussian distribution with mean $\mu^* = 0$, and standard deviation $\sigma^*$.

Then, the procedure outlined previously was repeated using simulated data sets, resulting in a new solubility prediction. This process was repeated 1000 times for each excess Gibbs free energy model, giving the predictions presented in Figure 6a. To investigate the effect of uncertainty magnitude, we perform sensitivity analysis at three levels for each thermodynamic quantity; $\sigma^*, 2\sigma^*$, and $5\sigma^*$, where $\sigma^*$ is defined as 0.01, 0.01, and 0.0001 for $\ln \gamma_p$, $\Delta H^{\text{dil}}$, and $\gamma_{p,\text{soln}}$ respectively. It should be noted that the chosen $\sigma^*$ values correspond to approximate uncertainty observed in aggregated data sets shown in Figures 2b, 3a, and 4b. From the available data, measurements of $\gamma_{p,\text{soln}}$ are more precise, in terms of % error, when compared to other thermodynamic quantities. Uncertainty levels translated to percentage errors are provided in Table 2. Resulting fits to experimental data are presented in the Supporting Information.

From Figure 6, solubility predictions for α-glycine derived from both activity coefficient models are somewhat insensitive to measurement uncertainty (at the levels investigated) between 273–340 K, as illustrated by thin blue and gray bands. However, as the prediction moves further from the reference temperature (298.15 K), prediction uncertainty increases as indicated by a spreading of the bands for all

Figure 5. Temperature-dependent α-glycine solubility predictions for various data modeling approaches. (Gray lines: 0th-order predictions. Blue lines: 1st-order predictions. Dashed lines: Scatchard–Hildebrand predictions. Solid lines: Scatchard–Hildebrand–Flory–Huggins predictions. Black open symbols: Direct measurements labeled as −predictions. Solid lines: Scatchard data is limited above 340 K, the approaches show good agreement in terms of solubility; −predictions. Blue lines: 1st-order predictions. Dashed lines: Scatchard variances data modeling approaches. (Gray lines: 0th-order predictions.)

https://doi.org/10.1021/acs.cgd.1c01217
Cryst. Growth Des. 2022, 22, 1691–1706

levels. As the level of uncertainty used to “blur” data increases, the uncertainty of predictions far from the reference temperature increases, which is what we expect. It is interesting to note that, even accounting for measurement uncertainty, predictions based on the different excess Gibbs free energy models are significantly different at temperatures far from the reference for predictions based $\sigma^*$ and $2\sigma^*$ uncertainty levels (Figure 6a and 6b). This suggests that model selection has an impact, independent of measurement uncertainty, on the validity of the method and motivates future work on finding the most appropriate model to collate consistent thermodynamic quantities of liquid solutions. However, at the $5\sigma^*$ level, we note that prediction bands overlap for the entire temperature range analyzed. From Figure 6d, we note the percentage error for each uncertainty level taken with respect to the best fit solubility predictions for each excess Gibbs free energy model. As shown, the percentage error in solubility increases as the prediction moves from the reference temperature, while at 400 K, the percentage error in solubility prediction is approximately 1, 3, and 8−10% for $\sigma^*$, $2\sigma^*$, and $5\sigma^*$, respectively.

The sensitivity of solubility predictions (based on the Scatchard–Hildebrand excess Gibbs free energy model) to independent thermodynamic data is shown in Figure 7. We

| quantity   | error (%) |
|------------|-----------|
| $\sigma^*$ | $\ln \gamma_w$ | 1 |
|            | $\Delta H_{\text{H}}$ | 1 |
|            | $c_{\text{p,soln}}$ | 0.01 |
| $2\sigma^*$| $\ln \gamma_w$ | 2 |
|            | $\Delta H_{\text{H}}$ | 2 |
|            | $c_{\text{p,soln}}$ | 0.02 |
| $5\sigma^*$| $\ln \gamma_w$ | 5 |
|            | $\Delta H_{\text{H}}$ | 5 |
|            | $c_{\text{p,soln}}$ | 0.05 |

Table 2. Levels Used in $\alpha$-Glycine Sensitivity Analysis

Figure 6. (a−c) Solubility predictions based on excess Gibbs free energy model parameters fit to “simulated” data sets. The gray fan represents the range of predictions for the Scatchard–Hildebrand model, and the blue fan represents the range of predictions of the Scatchard–Hildebrand–Flory–Huggins model. (d) Prediction uncertainty based on fan horizontal width. Gray refers to the Scatchard–Hildebrand model, and blue refers to the Scatchard–Hildebrand–Flory–Huggins model. The open black circles are selected direct measurements labeled as $\alpha$-glycine$^{25,26,28−30}$ and undefined glycine$^{9,33−39}$.
If we consider that the level of uncertainty in where a greater impact corresponds to larger percentage error. Δc

Glycine Reference Solubility

"note that, for the case where uncertainty in data is estimated by data points and generating new model fits, solubility predictions far from the reference temperature are impacted by

Figure 7. Uncertainty of the predicted α-glycine solubility based on the Scatchard–Hildebrand model with uncertainty of σx (gray), 2σx (blue), and 5σx (red) in the value of ln γw (solid lines), ΔHfl/n (dashed lines), and c[p,soln] (dashed-dotted lines).

If we consider that the level of uncertainty in c[p,soln] is 100 times lower than ΔHfl and ln γw, it suggests that, for accurate solubility predictions, precise measurements of c[p,soln] are required (i.e., ±0.1 J mol⁻¹ K⁻¹). However, it should be noted that, for the same level of uncertainty, both c[p,soln] and ΔHfl/γ have a similar impact, while for all levels, uncertainty in ln γw has comparatively lower impact. For example, at the 5σx level, the percentage uncertainty in predicted solubility at 400 K is 2.5%, 4%, and 5% for ln γw, ΔHfl, and c[p,soln] respectively.

β and γ-Glycine Solubility Predictions. In the context of the modeling framework presented for α-glycine, there is scope to predict the solubility of β and γ-glycine, given that direct measurements for ΔHfl and c′E are available (Table 1), alongside regressed excess Gibbs free energy models describing the thermodynamic properties of glycine–water solutions. However, for β-glycine, there is only one (unreliable) report of solubility data available, while for γ-glycine the solubility data is scattered (Figure 1). As discussed, predictions require an estimate for the solubility at the chosen reference temperature around which the expansion is based. While the solubility of α-glycine is accurately known, various data was collated to estimate β and γ-glycine solubility predictions as summarized in Table 3.

Table 3. Summary of Data Used to Estimate β- and γ-Glycine Reference Solubility

|       | β                 | γ                 |
|-------|-------------------|-------------------|
| ΔHfl(298.15 K) | (cβ/cα)(310 K)    | (cγ/cα)(310 K)    |
| Tβ and xβ   | xγ               | xγ               |
| γ/γ              | Tγ               |

In Table 3, Δµ°γ/α is an estimate of the polymorph free energy difference at 298.15 K, (cβ/cα)(310 K) is polymorph solubility ratios in water-antisolvent systems with various solvent compositions at 310 K, Tβ and xβ are eutectic temperature and composition measurements, xγ corresponds to direct solubility measurements of γ-glycine, xγ/xα corresponds to α/γ solubility ratios derived from direct solubility measurements reported together and Tγ/α corresponds to estimates for the temperature at which the relative stability of α and γ-glycine changes.

Eutectic Temperature and Composition. The eutectic temperatures for α-, β-, and γ-glycine in water have been reported by various authors and recently reviewed. From the available data (see Supporting Information (Eutectic Temperature and Composition)), we note that β-glycine has the most reported measurements, followed by γ and finally α with a single measurement.

The eutectic temperatures for β- and γ-glycine are consistent when the reported error is considered, while those reported for α-glycine are conflicting. We assume the reported eutectic temperature of ~3.6 °C for α is incorrect on the basis that it is the same as the reported value for β-glycine, and our α-glycine solubility predictions indicate ~2.8 °C is a more reasonable value. The expected eutectic temperature for each polymorph was estimated by taking the mean of the reported values and found to be ~2.8, ~3.7, and ~2.8 °C for α, β, and γ, respectively, on which we estimate an uncertainty of ±0.1 °C. The resulting eutectic temperatures are presented alongside selected freezing point measurements in Figure 8.

Figure 8. Expected eutectic temperatures and freezing curve The open black symbols are freezing curve measurements, and the gray dashed line is the best fit line. The solid lines are the estimated eutectic temperature, and the opaque bands indicate the uncertainty.

Blue refers to α-glycine, gray refers to γ-glycine, and red refers to β-glycine.

By evaluating the solution composition at which the eutectic temperature range for β- and γ-glycine intersects the freezing point line, we estimate their eutectic composition; giving xβ/α = 0.040 ± 0.001 and xγ/γ = 0.030 ± 0.001 for β and γ respectively. In the absence of reliable estimates for the solubility of β and γ-glycine at 298.15 K, the eutectic temperature and composition can be used as a point from which we can base our solubility predictions. For example, we can iteratively perform solubility predictions with different reference values and assess the resulting performance in relation to

https://doi.org/10.1021/acs.cgd.1c01217
Cryts. Growth Des. 2022, 22, 1691–1706
to the estimated eutectic temperature and composition. This would work for β-glycine; however, given that the eutectic temperature and composition estimate for γ is the same as α, the utility of the approach is limited for γ-glycine.

**Bouchard Ratios.** Bouchard et al. report measurements of α-, β-, and γ-glycine solubility in water-antisolvent mixtures of varying solvent composition at 310 K. Measurements for α- and γ-glycine are given in pure water, giving a solubility ratio of \( \frac{x_\gamma}{x_\alpha} = 0.974 \pm 0.012 \). Measurements of β-glycine in pure water are not reported due to rapid recrystallization to α-glycine. However, using available solubility measurement in antisolvent mixtures, we can estimate the solubility ratio of β and α in pure water by extrapolating the solubility ratio from known antisolvent mixtures to pure water (see Supporting Information). We find that this approach gives a reasonable estimate for \( \gamma/\alpha \) thus apply it to \( \beta/\gamma \). The estimated solubility ratio was found to be \( x_\beta/x_\gamma = 1.13 \pm 0.08 \).

**Polymorph Free Energy Difference Estimates.** The free energy differences for glycine polymorphs (α and β) with respect to γ-glycine at 298.15 K have been estimated from pure crystal heat measurements at temperatures between 0 and 298.15 K and enthalpy of solution measurements.22,63,70 The free energy differences are reported as \( \Delta \mu^{\alpha/\beta} = 157 \pm 145 \) J mol\(^{-1}\) K\(^{-1}\) and \( \Delta \mu^{\beta/\gamma} = 277 \pm 44 \) J mol\(^{-1}\) K\(^{-1}\), giving \( \Delta \mu^{\alpha/\beta} = 120 \) J mol\(^{-1}\) K\(^{-1}\). It should be noted that the reported uncertainty for \( \Delta \mu^{\alpha/\beta} \) is of the same magnitude as the difference; while no uncertainty was reported for \( \Delta \mu^{\beta/\gamma} \). From thermodynamics, it can be shown that the solubility of two polymorphs at a given temperature is related to the free energy difference:

\[
\beta (\mu^{\gamma/\alpha} - \mu^{\beta/\gamma}) = \ln \left( \frac{x^{\gamma/\alpha}}{x^{\beta/\gamma}} \right) = \ln \left( \frac{x^{\gamma/\alpha}}{x^{\beta/\gamma}} \right)
\]

where we make the assumption that the activity coefficient is approximately the same given that we expect the mole fraction solubility of each polymorph to be reasonably close. From eq 43 we estimate the solubility ratios to be \( x_\gamma/x_\alpha = 0.94 \pm 0.05 \) and \( x_\beta/x_\gamma = 1.05 \). It should be noted that the mole fraction solubility ratio estimated for \( \gamma/\alpha \) is in good agreement with those derived from direct measurements (Supporting Information).

**Solubility Predictions.** On the basis of the discussion above, solubility predictions for β and γ-glycine were developed based on the following. For β-glycine, the solubility prediction was fixed on the expected eutectic temperature of −3.7 °C using an iterative approach to find the “best” reference solubility at \( T_0 = 298.15 \) K, which was found to be 0.074 and 0.073 for the Scatchard−Hildebrand and Scatchard−Hildebrand−Flory−Huggins solution models, respectively. For γ-glycine, the ratio estimated by the free energy was used, giving \( x_{\gamma/298.15} = 0.052 \) (based on \( x_{\gamma/298.15} = 0.056 \)). The resulting solubility ratio, and eutectic point estimates are presented in Figures 9–11, respectively.

As expected, the choice of the excess Gibbs free energy model has a similar impact on solubility predictions for β and γ-glycine, as was seen for α-glycine, resulting in diverging solubility predictions away from the reference temperature. As shown in Figure 9, the effect of excess Gibbs free energy model on solubility predictions for γ-glycine is similar to α-glycine—the predictions begin to diverge around 340 K. However, it is interesting to note predictions for β-glycine diverge much closer to the reference temperature (i.e., 298.15 K). This is likely a result of β-glycine having a much higher solubility at 298.15 K, which, as shown in Figure 2, corresponds to diverging estimates for \( \partial \ln \gamma/\partial x_\gamma \) and \( \partial^2 \ln \gamma/\partial x_\gamma^2 \), for each excess Gibbs free energy model. As such, we suggest 273.15–310 K as reliable range for our β-glycine solubility predictions. However, we expect reliability could be improved with accurate water activity data solution compositions above 0.055 glycine mole fraction. This further highlights that prediction accuracy of the presented approach can be improved through more accurate estimation of partial derivatives of relevant thermodynamic quantities.

In addition, the mole fraction solubility ratios (across a wide range of temperature) are similarly impacted. It is interesting...
Sensitivity analysis was used to assess accuracy of solubility predictions with respect to underlying measurement uncertainty, as well as underlying excess free energy models used to derive required thermodynamic quantities. This provided a range of plausible solubility predictions far from the chosen reference temperature, and in particular above 340 K, where there is lack of reliable solubility data for glycine. Finally, we applied the approach to $\beta$ and $\gamma$-glycine where previous solubility data is inconsistent or limited, providing estimates for their aqueous solubility between 273–310 and 273–330 K, respectively.

The approach introduced here provides a novel framework for how various thermodynamic data can be used in concert to predict the temperature dependent solubility of crystal polymorphs. This will be useful for systems where direct measurements of solubility are challenging for one or more polymorphs, and compounds that undergo thermal decomposition or polymorph transition prior to melting.

ASSOCIATED CONTENT

Supporting Information

The following files are available free of charge. The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.cgd.1c01217.

- Solubility data fits, freezing point data fits, model parameter definitions, supplementary sensitivity analysis figures, additional derivations, and tabulated solubility predictions (PDF)

AUTHOR INFORMATION

Corresponding Author

Leo Lue — Department of Chemical and Process Engineering, University of Strathclyde, Glasgow G1 1XJ, U.K.; orcid.org/0000-0002-4826-5337; Email: leolue@strath.ac.uk

Authors

Andrew Manson — Department of Chemical and Process Engineering, University of Strathclyde, Glasgow G1 1XJ, U.K.; orcid.org/0000-0002-6785-8891
Jan Sefcik — Department of Chemical and Process Engineering, University of Strathclyde, Glasgow G1 1XJ, U.K.; EPSRC Continuous Manufacturing & Advanced Crystallisation (CMAC) Future Manufacturing Research Hub, University of Strathclyde, Glasgow G1 1RD, U.K.

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.cgd.1c01217

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to acknowledge support for A.M. from the Carnegie Trust for the Universities of Scotland (Grant PHD008559). We would also like to acknowledge support for J.S. from EPSRC Continuous Manufacturing and Advanced Crystallization Future Manufacturing Research Hub (Grant EP/P006965/1).

Figure 11. Glycine polymorph mole fraction solubility ratio predictions based on eq 16, with coefficients derived from different glycine–water solution models. Blue refers to the $\gamma$–$\alpha$ ratio, and red refers to $\beta$–$\alpha$ ratio. The dashed lines is the prediction of the Scatchard–Hildebrand model, and the solid lines Scatchard–Hildebrand–Flory–Huggins model. The open symbols are ratios derived from direct measurements (19–27) (see Supporting Information), and the pluses are estimates from free energy ratios, stars are estimates from eutectic temperatures and freezing point data, triangles are estimates from ref 6, and gray squares are $\alpha$–$\gamma$ crossover estimates.2,6,1,17. Bars indicate uncertainty estimates.

to note (as shown in Figure 11) the Scatchard–Hildebrand–Flory–Huggins model correctly predicts the $\gamma$/$\alpha$ crossover in the range of temperatures reported in literature; however, the Scatchard–Hildebrand model fails to do so, despite coming very close. This could be interpreted as an indication that the composition dependence of the water activity coefficient is better described by the Scatchard–Hildebrand–Flory–Huggins model. Both solution models predict $\beta$-glycine to be least stable (i.e., highest solubility) in the temperature range explored, agreeing with qualitative observations in the literature. However, predictions for $\beta$/$\alpha$ based on the expected eutectic temperature appear to be inconsistent with independent ratio estimates derived from antisolvent mixture solubility data and free energy difference estimates, though this could be attributed to significant uncertainty associated with both estimates.

Finally, we note that solubility predictions based on different solution models are approximately identical in the eutectic region (as shown in Figure 10), where the predicted eutectic temperature and composition for $\alpha$-glycine agrees extremely well with data available in literature, as does $\gamma$ (noting that models for $\alpha$ and $\gamma$ were not based on the eutectic temperature at all). Our predictions suggest that $-2.7$ °C is an appropriate eutectic temperature for $\gamma$-glycine.

CONCLUSIONS

A novel approach to estimate the temperature dependence of solubility was presented that relies on experimentally accessible thermodynamic data at a single temperature. The approach was applied the glycine–water system, and it was found that, between 273–340 K, the approach provided a solubility prediction that agreed very well with available direct solubility measurements for $\alpha$-glycine.
(43) Ellerton, H. D.; Reinfields, G.; Mulchay, D. E.; Dunlop, P. J. Activity, Density, and Relative Viscosity Data for Several Amino Acids, Lactamide, and Raffinose in Aqueous Solution At 25 °C. J. Phys. Chem. 1964, 68, 398–402.

(44) Kurhe, D. N.; Dagade, D. H.; Jadhav, J. P.; Govindwar, S. P.; Patil, K. J. Studies of Enthalpy-Entropy Compensation, Partial Entropies, and Kirkwood-Buff Integrals for Aqueous Solutions of Glycine, L-Leucine, and Glycglycine At 298.15 K. J. Phys. Chem. B 2009, 113, 16612–16621.

(45) Smith, E. R.; Smith, P. K. The Activity of Glycine in Aqueous Solution At Twenty-Five Degrees. J. Biol. Chem. 1937, 117, 209–216.

(46) Lilley, T. H.; Scott, R. P. Aqueous Solutions Containing Amino-Acids and Peptides. Part 2. Gibb’s Function and Enthalphy Behaviour of the Systems Urea + Glycine, Urea + α-alanine, Urea + α-aminoanmonic Acid and Urea + Glycglycine At 298.15 K. J. Chem. Soc., Faraday Trans. 1 1976, 72, 184.

(47) Salabat, A.; Neshat, S.; Fazlali, A. Activity Coefficients of Glycine, α-Alanine and α-Lyvaline in Aqueous Solutions Containing MgSO₄, At 298.15 K; Experimental Determination and Correlation. Fluid Phase Equilib. 2012, 314, 198–202.

(48) Pinho, S. P. Water Activity in Aqueous Amino Acid Solutions, With and Without KCl, At 298.15 K. J. Chem. Eng. Data 2008, 53, 180–184.

(49) Ninni, L.; Meirelles, A. Water Activity, pH and Density of Aqueous Amino Acid Solutions. Biotechnol. Prog. 2001, 17, 703–711.

(50) Sadeghi, R.; Gholamireza, A. Thermodynamics of the Ternary Systems: (water + Glycine, α-Alanine and α-Serine + Di-Ammonium Hydrogen citrate) From Volumetric, Compressibility, and (vapour + liquid) Equilibria Measurements. J. Chem. Thermodyn. 2011, 43, 200–215.

(51) Sturtevant, J. M. The Heats of Dilution of Aqueous Solutions of Glycine At 25 °C. J. Am. Chem. Soc. 1940, 62, 1879–1879.

(52) Humphrey, R.; Hedwig, G.; Watson, I.; Malcolm, G. The Partial Molar Enthalpies in Aqueous Solution of Some Amino Acids With Polar and Non-Polar Side Chains. J. Chem. Thermodyn. 1980, 12, 595–603.

(53) Gucker, F. T.; Pickard, H. B.; Ford, W. L. The Heats of Dilution of Aqueous Solutions of Glycine and Glycolamide, and Other Thermodynamic Properties of Glycine at 25°C. J. Am. Chem. Soc. 1940, 62, 2698–2704.

(54) Wang, X.; Xu, L.; Lin, R.; Sun, D. Enthalpies of Dilution of Glycine, L-Alanine and L-Serine in Aqueous Potassium Chloride Solutions. Thermochim. Acta 2005, 425, 31–37.

(55) Wallace, W. E.; Offutt, W. F.; Robinson, A. L. The Heats of Dilution of Aqueous Solutions of Glycine At 25°C. J. Am. Chem. Soc. 1943, 65, 347–350.

(56) Pálecz, B. Enthalpies of Solution and Dilution of Some L-α-Amino Acids in Water at 298.15 K. J. Therm. Anal. Calorim. 1998, 54, 257–263.

(57) Li, S.; Hu, X.; Lin, R.; Zong, H. Enthalpic Interaction of Glycine in Aqueous Glucose and Sucrose Solutions At 298.15 K. Thermochim. Acta 1999, 342, 1–6.

(58) Ren, X.; Ni, Y.; Lin, R. Enthalpies of Dilution of Glycine, L-Serine and L-Lysine in Mixtures of Water and N2,N-Dimethylformamide At 298.15 K. Thermochim. Acta 2000, 348, 19–24.

(59) Gucker, F. T.; Allen, T. W. The Densities and Specific Heats of Aqueous Solutions of dl-α-Alanine, β-Alanine and Lactamidol2. J. Am. Chem. Soc. 1942, 64, 191–199.

(60) Downes, C. J.; Hakin, A. W.; Hedwig, G. R. The partial molar heat capacities of glycine and glycyglycin in aqueous solution at elevated temperatures and at p = 10.0 MPa. J. Chem. Thermodyn. 2001, 33, 873–880.

(61) Spink, C.; Wadsö, I. Thermochemistry of solutions of biochemical model compounds. 4. The partial molar heat capacities of some amino acids in aqueous solution. J. Chem. Thermodyn. 1975, 7, 561–572.