An Evaluation of Methods for Estimating the Solubility of Solids in Liquids

by

Andrew C. Manson

MEng, University of Strathclyde (2019)

Submitted to the

Department of Chemical and Process Engineering

in partial fulfillment of the requirments for the degree of

Doctor of Philosophy in Chemical Engineering

at the

University of Strathclyde

July 2023

Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyrights Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

Abstract

Crystallisation is an example of a separation and purification technique used ubiquitously in chemical manufacturing processes. For the design, development and optimisation of efficient crystallisation processes, accurate solid-liquid equilibria (i.e., solubility) is necessary. Solubility is directly measurable; however, despite years of development, available techniques have drawbacks related to the time required to complete, the amount of material required, economic cost, as well as technical challenges under certain conditions (i.e., solid form, temperature, pressure and composition).

In this thesis, new model-based methods rooted in classical equilibrium thermodynamics are developed that allow the solubility of a solid in a liquid to be estimated as a function of temperature. Crucially, these methods are developed in a manner that reduce reliance on direct solubility measurements to a single temperature, and make it possible to predict solubility through measurement and mathematical modelling of thermodynamic quantities other than solubility.

First, an approximate version of the van't Hoff equation is developed and used to predict the aqueous solubility of glycine polymorphs using thermodynamic data reported in the literature at 298.15 K, only. For α-glycine, predictions in excellent agreement with direct solubility data from 260–340 K were made and additional predictions for β and γ-glycine, where less certain solubility data are available in the literature, are also provided.

Then, an analytical solubility model is proposed by interpreting a criteria for thermodynamic equilibrium (i.e., equal chemical potential) in terms of a second-order Taylor series expansion. Again, this approach makes it possible to predict solubility using only a single solubility measurement and a collection of thermodynamic quantities evaluated at the chosen reference temperature. Encouraging results are presented for the aqueous solubility of α-glycine, sucrose and urea—all of which have a meaningful amount of directly determined solubility, as well as thermodynamic quantities measured at 298.15 K.

In the final part of this thesis, we illustrate how to parameterise Gibbs free energy models in such a manner that solubility can be estimated precisely as a function of temperature by combining only a single measure of solubilty and a broad range of thermodynamic quantities. This retains the key advantages of the approximate approaches mentioned already, but makes it possible to reliably estimate solubility across an unbounded range of temperature, in principle. As an illustrative

example, we use thermodynamic data reported in the literature for the sucrose-water system to make predictions, noting it is possible to generate estimations in excellent agreement with directly measured solubility data from the freezing point to the assumed solute-solution-vapor triple point.

In summary, this thesis helps to fill a gap left by existing approaches in the ever growing suite of methods at the disposal of practitioners looking to estimate the solubility of solids in liquids.

Acknowledgements

It is my pleasure to acknowledge those who invested their time and energy supporting this endeavor.

First, I'd like to recognise the contributions of my supervisors Dr Leo Lue and Prof. Jan Sefcik. Leo and Jan—thank you. I learned a great deal under your guidance and I'm very grateful for the opportunity to have worked with you both.

Second, the work described in this thesis would not have been possible without the support of the Carnegie Trust for the Universities of Scotland. To all those who represent the Trust, I'd like express my sincere gratitude for granting me this opportunity.

Third, I'd like to acknowledge those who formed part of the Chemcial and Process Engineering Department at the University of Strathclyde at one point or another during my time as a research student for creating a stimulating work (and social) environment.

Finally, I'd like to give thanks to everyone in my life (past, present and future)—you make it all worthwhile.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Chemical manufacturing processes facilitate transformation of raw materials into useful end products that add value to our lives. Typically, multi-component mixtures are encountered during chemical manufacturing processes and, given that the primary goal of any manufacturing process is to produce products with a specific chemical nature and purity, having techniques to effectively separate and purify often complex multi-component mixtures is essential. Furthermore, the importance of effective separation and purification unit operations extends beyond the primary goal of producing desired end products, with an increasing focus on limiting waste and enhancing the recovery of potentially valuable secondary products.

A broad range of techniques exist allowing separation of products, by-products, unused raw materials and waste streams [1], including: distillation, absorption, and extraction [2]. Crystallisation, where chemical species existing in (usually, but not necessarily) multi-component liquid mixtures aggregate forming separate crystalline phases when conditions allow, is another example. As described in [3], crystallisation has been used in chemical engineering processes for many years. For example, it is thought that 'early civilizations' crystallized salt from sea water through evaporation. Then, crystallizers were shallow ponds, rather than sophisticated apparatus seen in contemporary chemical manufacturing facilities.

Crystallisation is attractive for industrial applications where the goal is to produce solid products free from impurities because it combines solid formation and purification in a single step [4]. More specifically, through rational design of a crystallisation process, it is possible to produce 'solid particles with desired size, shape, crystal form and chemical purity in a reproducible manner', which is crucial as all of the aforementioned properties can affect the physio-chemical properties of produced solid phases [4]. In certain applications, it is also thought that crystallisation is generally less energy intensive compared to alternative separation and purification techniques [2], due to the mild temperature and pressures used.

Solubility is a fundamental property governing crystallisation processes, setting a theoretical

limit on the amount different 'solutes' can exist in a liquid solution at thermodynamic equilibrium. Solubility varies with temperature, pressure and solution composition; as such, accurate solubility data is required for the design, development and optimisation of robust industrial crystallisation processes. Many techniques exist allowing estimation of solubility through direct experimental methods and this approach is widely applied in practice. However, although seemingly simple to perform, direct determination of solubility can be challenging for reasons relating to crystallisation phenomena (e.g., controlling crystal form and equilibration time), as well as technical challenges (e.g., temperature control, mass transfer, etc). Furthermore, a common critique of experimental approaches is that they are expensive in terms of time, material and economic cost. As such, there is motivation to develop approaches beyond direct measurement that enable reliable estimation.

Phase equilibria (e.g., solubility) can be described from the Laws of Thermodynamics. Not only does this make it possible to construct models describing the dependence of solubility on process parameters that can be used to correlate direct solubility measurements, it also makes it possible to construct models where solubility is expressed in terms of experimentally measurable thermodynamic properties other than solubility. As such, approaches that work alongside, or in place of, purely experimental approaches are possible.

At a minimum, estimating solubility through equilibrium thermodynamics requires mathematical representations for the Gibbs free energy of the pure crystalline solute, as well as the multi-component liquid solution, across the entire range of temperature, pressure and composition that solubility is required. Broadly speaking, the majority of approaches developed within this framework and applied in practice are based on interpreting these functions with reference to the pure solute melting properties (e.g., enthalpy of melting and melting temperature). Doing so, it is theoretically possible to 'predict' solubility across the entire range of temperature for a given system without ever knowing the solubility, which is highly attractive given known challenges with approaches that rely on direct measurements. However, this widely applied theoretical approach has limited utility for solutes with melting properties that are difficult to measure due to chemical degradation prior to melting or transformation into a different solid form.

In this thesis, the application of equilibrium thermodynamics to the general problem of estimating the solubility of solids in liquid is explored and, more specifically, we focus on theoretical descriptions of binary systems comprising single-component solutes and solvents. The primary objective is to evaluate new approaches that allow the solubility of solids in liquids to be estimated as a function of temperature in manner that retain the predictive nature of established model-based approaches but without potential shortcomings introduced through reliance on measurement of solute melting properties or significant amounts of directly measured solubility data. Integral to methods evaluated in this thesis is a change of perspective in the interpretation of underlying Gibbs free energy functions of phases, by referencing to a specific temperature and composition other

than the pure solute melting point. In Chapter 3 and 4, this is used to develop locally accurate, predictive models for temperature dependent solubility, and a global approach in Chapter 5.

The remainder of this thesis is structured as follows. In Chapter 2, relevant background required to place original contributions described in this thesis is presented. First, the fundamentals of crystallisation are introduced, including nucleation, crystal growth and solubility, followed by a survey of experimental approaches for the direct determination of solubility, motivating model-based approaches. Then, the fundamentals of classical equilibrium thermodynamics are introduced through a series of postulates originally formulated in [5]. This leads to a thermodynamic definition of phase equilibrium, after which the concept of a thermodynamic potential and the Gibbs free energy are introduced. This chapter is completed with a review of approaches for the correlation and prediction of solubility based in the equilibrium thermodynamics.

In Chapter 3, a method based on the van't Hoff equation applied to solid-liquid equilibrium which allows extrapolation of temperature dependent solubility around a chosen reference temperature based on classical thermodynamic quantities is presented. Contrary to traditional approaches, the melting properties of the pure solute are not required to make predictions, making it applicable to systems where melting properties are challenging to measure. In principle, this approach requires only a single solubility measurement at the chosen reference temperature, reducing reliance on potentially challenging solubility measurements. This approach is benchmarked by estimating the aqueous solubility of α-glycine, where predictions in good agreement with measurements reported in literature around 298.15 K, are obseverved. The sensitivity of the approach to property estimation is examined through a Monte Carlo approach and additional predictions for the aqueous solubility of β and γ-glycine are presented, both of which have limited experimental studies present in the literature.

In Chapter 4, the approach described in Chapter 3 is refined. This leads to an analytical model for solubility as a function of temperature expressed in macroscopic thermodynamic quantities. Further, this approach retains one of the key characteristics of the approach described in Chapter 3; no reliance on the pure solute melting properties at the normal melting point. Extrapolations for two new binary systems—sucrose and urea (both in water)—with more complete solubility data (when compared to the glycine-water system described in Chapter 3) are presented. As with the approach in Chapter 3, predictions are in good agreement with available data and model-based sensitivities examined.

In Chapter 5, one of the main ideas from Chapter 3 and 4 (i.e., estimating solubility based on a single point on the solubility curve) is extended to a more general modelling framework, free from the approximations introduced in the approaches presented in those chapters. For a specific approach to single and multi-component Gibbs free energy model construction, it is shown that global descriptions are possible with only a single solubility measurement, where pure

crystal melting properties are not required. The general framework is explored in the context of the sucrose-water system and solubility predictions provided. As shown, for a specific modelling interpretation of pure sucrose crystal and sucrose-water liquid mixture, it is possible to predict the sucrose aqueous solubility in excellent agreement with directly measured data, across the entire range of temperature.

This thesis is closed with a chapter summarising the main conclusions of the work, as well as recommendations for future work.

Chapter 2

Background

At the core of this thesis is the problem of estimating the solubility of solids in liquids, relevant in the broader context of the design, development and optimisation of industrial crystallisation processes. In this chapter, relevant background for the work described in the remainder of this thesis is provided and is structured as follows. First, the fundamentals of crystallisation are described, including: nucleation, crystal growth and solubility. With the importance of solubility as necessary component for the design of efficient crystallisation processes elucidated, established approaches used to estimate solubility are discussed starting with a summary of direct experimental measurements techniques (and associated challenges) and followed by model-based approaches. Of the model-based approaches introduced, those derived from the Laws of equilibrium thermodynamics are the focus of this thesis. As such, fundamentals of equilibrium thermodynamic theory are subsequently described, covering the conditions for phase equilibria, as well as the introduction of key thermodynamic quantities such as the chemical potential and Gibbs free energy, both of which are used extensively in later chapters. The discussion on thermodynamics is completed by introducing the role of mathematical modelling as a tool to connect thermodynamic theory and experiments. The final part of this chapter is focused on surveying progress in the general area of solubility estimation via equilibrium thermodynamics and motivating the work presented in the remainder of this thesis. More specifically, the development of approaches anchored on the melting properties of solutes is reviewed.

2.1 Crystallisation

As described in Chapter 1, the aim of this thesis is to evaluate methods to estimate the solubility of solids in liquids using the classical equilibrium thermodynamics as a framework to develop mathematical descriptions of solubility in the context of industrial crystallisation. More specifically, we focus on crystallisation from liquid solutions. Throughout the remainder of this thesis, we adopt

the standard terminology where the crystallising component is referred to as the 'solute' and the remaining components are termed 'solvents'.

Briefly, a crystal is a characterised by a regular microscopic structure (i.e., lattice) related of the chemical nature of the particles (i.e., molecules, atoms or ions) that it comprises [6]. It is possible for crystals formed by particles of the same chemical nature to form different crystal structures called 'polymorphs' (c.f. polymorphism), characterised by different lattice energies. Although not dealt with explicitly in the work described in this thesis, for completeness, it should be noted that in addition to polymorphism, it is possible for a crystal phase to comprise multiple chemically distinct particles, and common examples include: hydrates, solvates, solid solutions and co-crystals. In certain situations, multi-component crystals are advantageous due to enhanced physio-chemical properties; however, may also present challenges. It should be noted that several texts are available presenting more comprehensive discussions on crystallisation and interested readers should refer to these where relevant [3, 6, 7, 8].

2.1.1 Nucleation and Crystal Growth

Crystallization (of a solid from a liquid) is a phenomenon characterised by the transfer of substances that previously formed part of a homogeneous liquid solution to separate crystalline phases. This process is assumed to consist of two stages—nucleation and crystal growth. Crystal growth is the process by which an already existing stable crystalline phase increases in size through aggregation of material from an adjacent homogeneous liquid solution. There are various theorectical interpretations describing crystal growth and interested readers should refer to [9] for further details.

Nucleation is defined as the process through which crystal 'nuclei' are formed. It is accepted that nucleation occurs via different mechanisms, namely primary and secondary nucleation. Primary nucleation corresponds to the spontaneous formation of a stable nuclei in the absence of crystalline solid. Primary nucleation is further classified as being either 'homogeneous'—where nuclei appear in the bulk solution in the absence of another solid interface (which may promote nucleation), or 'heterogeneous'—where nuclei appear in the presence of a solid interface [3]. Like crystal growth, the fundamental mechanism of primary homogeneous nucleation is not yet understood due to the small time and length scales involved [10]. Initially, it was thought that crystal nucleation could be described in an analogous way to droplets from a vapor—giving a model typically referred to as 'Classical Nucleation Theory' (CNT) [4, 6, 8]. Classical Nucleation Theory describes the nucleation in terms of the thermodynamic Gibbs free energy (Section 2.2) by considering the difference between the free energy of a solid solute and a solute in solution [6], stated

mathematically as:

$$\Delta G = \frac{4}{3}\pi r^3 \Delta g_v + 4\pi r^2 \gamma, \tag{2.1}$$

where γ denotes the inter-facial tension between a solid solute particle and surrounding solution, Δg_v denotes the free energy change of solid-liquid transformation per unit volume and r denotes solid solute particle radius (noting that it is assumed to be spherical). In the context of CNT, a nucleus is stable when the radius is greater than or equal to the 'critical radius' (i.e., when $d\Delta G/dr = 0$), given by $-2\gamma/\Delta g_v$. However, failure of CNT to describe empirical observations, coupled with more recent computational and experimental results has led to additional theories. For more detailed discussion, readers should refer to [4] where the topic is discussed in greater detail.

Secondary nucleation corresponds to the formation of nuclei in the presence of already formed crystals and is viewed as a more readily controllable alternative to primary nucleation for industrial applications [10]. It is thought that there are different mechanisms that drive secondary nucleation, such as: initial breeding, dendritic or needle propagation, and crystal attrition [10]. Compared to primary nucleation, secondary nucleation can occur at lower supersaturation levels [8], and can be achieved by through introduction of so-called 'seed' crystals into supersaturated solutions.

2.1.2 Solubility and Supersaturation

For the crystallisation of solids from liquid mixtures, solubility refers to the total amount of the crystallising substance that will dissolve in a solvent solution at fixed temperature, pressure and solvent composition. For a given solute, solubility can vary significantly as processing parameters (e.g., temperature, pressure and solvent composition) change. Focusing on the effect of temperature, this is illustrated in Fig. 2.1 for sucrose, sodium chloride and urea in water. As shown, the dependence of solubility can vary dramatically for different solutes in the same solvent. Although the general trend is for solubility to increase with temperature (e.g., urea and sucrose in water), it is possible to have an extremely weak temperature dependence (e.g., sodium chloride in water), and in some cases to decrease with increasing temperature (referred to as retrograde solubility).

Focusing on the effect of solvent composition at fixed temperature and pressure, where the solvent could be pure, or a mixture of different compounds, this is illustrated in Fig. 2.2 for sucrose in water, ethanol and mixtures of both at 298.15 K. As shown, sucrose is practically insoluble in pure ethanol. However, as the solvent composition is varied to increasing proportions of water, the solubility of sucrose increases towards it's solubility in pure water.

In addition to temperature, pressure and solvent composition, crystal structure is another factor that may impact solubility. This is shown for the solubility of mefenamic acid (which has at least two polymorphs—I and II) in ethyl acetate between 298.15 and 333.15 K in Fig. 2.3. As shown,

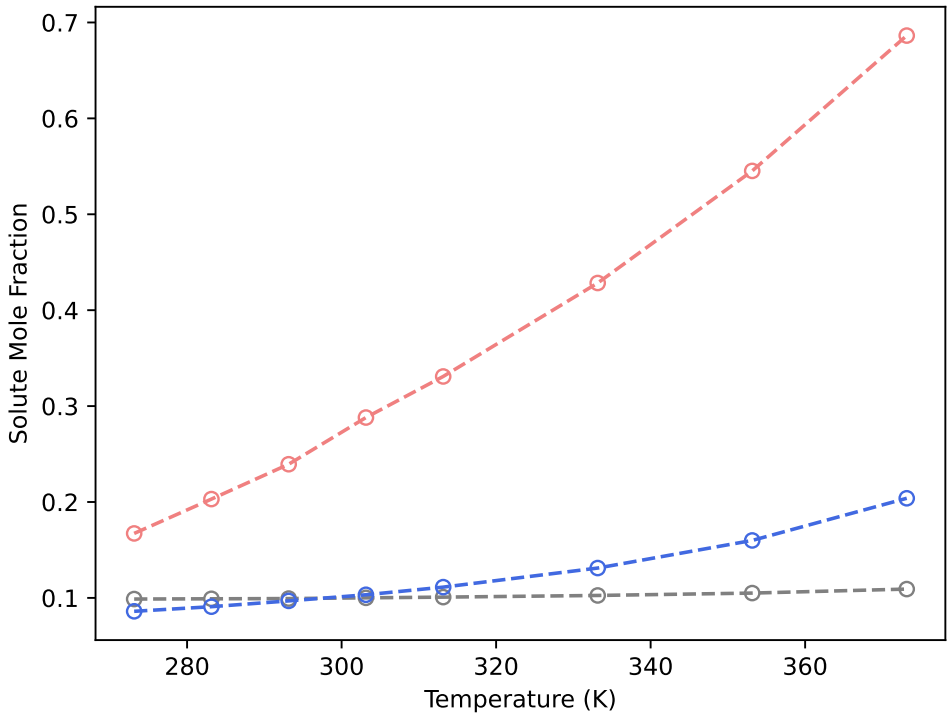


Figure 2.1: Aqueous solubility of sodium chloride (grey), sucrose (blue) and urea (red) between 273.15–373.15 K shown via open circles (data taken from [6]). Dashed lines given as a guide for the eye.

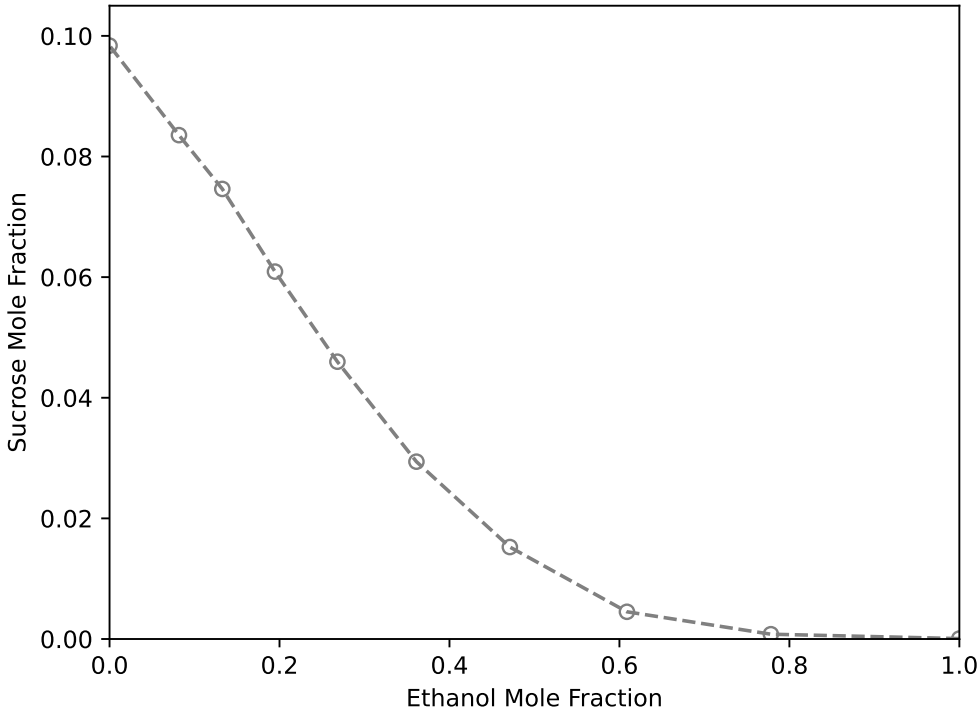


Figure 2.2: Solubility of sucrose in water, ethanol and mixtures of both at 298.15 K shown via open grey circles (data taken from [11]). Dashed line given as a guide for the eye.

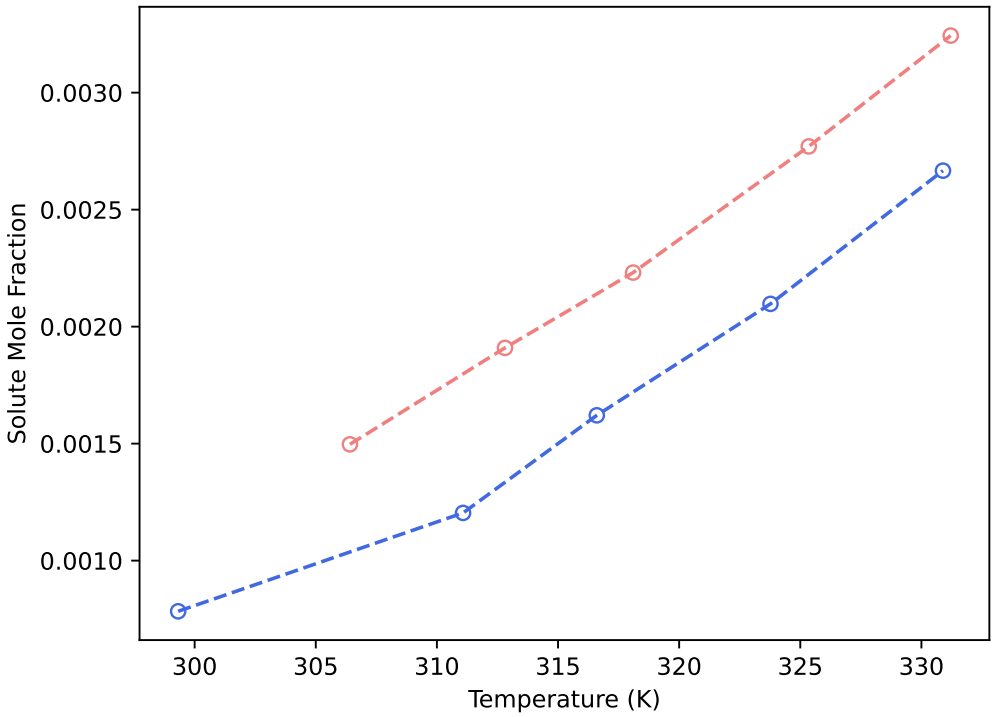


Figure 2.3: Solubility of mefenamic acid in ethyl acetate between 298.15–333.15 K shown via open circles (blue is form I, red is form II and data taken from [12]). Dashed lines given as guide for the eye.

form I is less soluble at all temperatures given, which has implications for the design, development and optimisation of crystallisation processes. As an aside, it should also be noted that polymorphs may have drastically different physiochemical properties besides solubility leading to potentially hazardous situations in a pharmaceutical context [13].

In the examples described above, it was assumed that all phases (i.e., solute and solvent) comprised particles of the expected chemical nature; however, systems encountered in practice are susceptible to impurities which is another factor that can have a material impact on solubility [3]. For example, as originally shown in [14] and highlighted in [15], presence of water impacted the solubility on an unspecified 'pharmaceutical development' compound's solubility in ethyl acetate. In [15], it was noted that additional water was present in the solute-solvent liquid mixture due to a previous processing step, and highlight potential challenges with experimental solubility determination in an environment where contamination is possible. For the specific case where water is present, this result is particularly relevant for hygroscopic solutes and/or solvents that may be exposed to water prior to any attempt to estimate solubility. It should be noted that this is not exclusive to water considering that mixing solvents is an approach to either increase or decrease solubility (referred to co-solvent and anti-solvent, respectively).

As noted, solubility sets a theoretical limit on the amount of solute that will dissolve in a fixed amount of solvent at fixed temperature, pressure and solvent composition. However, for crystallisation to occur, the amount of solute in the solvent must exceed the limit set by solubility at the corresponding conditions (i.e., temperature, pressure and solvent composition). When a liquid mixture contains a greater amount of solute than the associated solubility, the solution is often referred to as being 'supersaturated'.

Supersaturation is a necessary, but insufficient criterion for instant, spontaneous crystallisation. Experimental studies have shown that there is a so-called 'meta-stable zone' that beyond the solubility curve as illustrated in Fig. 2.4. In the meta-stable zone, the propensity for nucleation to occur is random and characterised by the induction time, which relates the period between achieving supersaturation and crystallisation being observed. Typically, supersaturation and induction time are related (i.e., increasing supersaturation decreases induction time). Beyond the meta-stable zone, the likelihood of instantaneous crystallisation increases drastically.

The metastable zone width is generally not viewed as a fundamental property of a solute-solvent system and varies with crystallisation conditions (i.e., scale, rate of supersaturation, etc) as described in [16]. Understanding where the metastable zone likely exists for a given crystallisation process is of interest because uncontrolled, spontaneous crystallisation at supersaturations greater than the metastable zone width may lead to crystals with irreproducible properties (i.e., size, shape, crystal form and chemical purity), which can make secondary processing (e.g., drying, tablet formation, etc.,) more challenging in the context where the product is to be delivered in

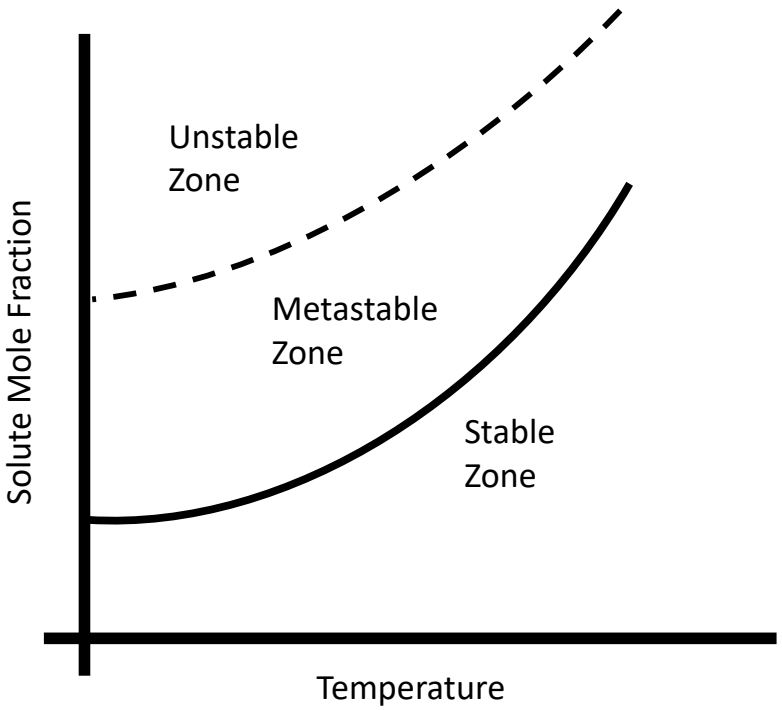


Figure 2.4: Solubility curve and meta-stable zone limit for binary solute-solvent system. Black solid line corresponds to solubility curve. Black dashed line corresponds to meta-stable zone limit.

solid form.

In practice, supersaturation is achieved in several ways [3, 17], including: cooling—where a solution at an initially high temperature is cooled, evaporation—where volatile solvent is evaporated, anti-solvent—where additional solvent (of different chemical nature) is added, as well as manipulating solution pH or by chemical reaction. For a given system, the method of attaining supersaturation and subsequent crystallisation ultimately depends on the relationship between solubility and various process parameters described above.

2.1.3 Estimating Solubility

Thus far, the importance of understanding how the solubility of a solute in a solvent varies with process parameters such as temperature and solvent composition for the design, development and optimisation of industrial crystallisation unit operations, given it's fundamental relationship to the crystallisation phenomenon, has been illustrated. In this section, methods used to estimate solubility are described starting with a description of direct experimental methods, followed by model-based approaches.

Direct Experimental Measurement

It is possible to measure solubility using experimental techniques and there are numerous approaches available. Based on current understanding of the crystallisation/dissolution process, regardless of the experimental approach taken to estimate solubility, key requirements for reliable estimation are:

1. attainment of thermodynamic equilibrium
2. characterisation of solid form and solution composition at equilibrium

As described in [18] and [19] and summarised in a recent editorial [20], one way to classify solubility measurement techniques is based on how saturation composition is estimated, leading to two broad classes of approaches.

The first are 'synthetic' (also known as 'polythermal' [6]) methods, where a known amount of solute is added to a container with a known amount of solvent. Then, temperature or pressure is varied until complete dissolution of the solid solute is observed either directly through visual techniques (e.g., turbidity), or detected via other means (e.g., vapor pressure, conductivity, heat flow, etc). For example, it has been shown that Differential Scanning Calorimetry (DSC) can be used to determine solubility indirectly by monitoring the thermal behavior of the dissolution process [21]. For the systems studied in [21], this approach was shown to give results in good agreement with equilibration-based (described shortly) techniques at moderate temperatures, as

well providing estimates for solubility at temperature close to (and above, which may be due to increased pressure in the crucible) the solvent boiling point (water). This approach was also extended to polymorphic systems [12]. Generally speaking, synthetic approaches are viewed as being faster than alternatives. However, given that crystallisation is a kinetic process, it is crucial that an appropriate rate of heating (for the case where solubility is inferred by complete dissolution of solid solute) is chosen to allow the system to achieve equilibrium.

Alternatively, there are 'analytical' (a.k.a equilibration) methods, where temperature and pressure are fixed while a system initially comprising an excess of solid solute and solution, or a previously supersaturated solution, are held in a container and allowed to reach thermodynamic equilibrium via dissolution or crystallisation, respectively. When an excess of solute is allowed to equilibrate with solution, this is sometimes referred to in the literature as the 'shake flask method', and quoted as being the 'gold standard' approach for solubility determination [13]. A key aspect of analytical approaches is the determination of equilibrium solution concentration, for which there are numerous approaches, including (but not limited to): gravimetric methods, high-performance liquid chromatography (HPLC), as well as nuclear magnetic resonance (NMR) and ultraviolet-visible (UV-vis) spectroscopy [15]. In contrast to 'synthetic' approaches described already, it has been shown that analytical approaches can take significant time to achieve equilibrium (inferred by imperceptible changes in solution composition as a function of time), with some systems requiring in excess of 48 hours [22], driven by the dissolution kinetics of the system under study, which has implications in practical settings where sufficient time to generate requisite solubility data may not be available. It should also be noted that challenges with temperature control during equilibration and solution sampling are additional technical challenges faced by this approach.

As might be expected, it has been suggested that no one measurement technique is universally applicable to all solid-liquid systems and the most appropriate choice largely depends on the context in which solubility data are required [3, 6], in terms of time, amount of material, reliability of results, etc. As an example, recent work assessing three experimental solubility estimation techniques (polythermal, gravimetric equilibration and solute addition) for their ability to provide reliable measurement of α-glycine solubility in water across a broad range of temperature indicated that each approach had relative advantages and drawbacks [23]. Another recent example highlighting possible system specific challenges is that reported in [24], where it was found that measuring the solubility of tolbutamide polymorphs (form I and II) in n-propanol using a gravimetric equilibration introduced significant differences (i.e., 5–15 g/kg, between 275 and 305 K) in solubility estimates when compared to those produced using a polythermal method due to poor settling properties of excess solid, impacting the authors ability to accurately characterise the equilibrium solution concentration. Interestingly, the issue was specific to only one polymorph (i.e., form II), but serves to highlight the unpredictable nature of solute-solvent systems and the practical challenges with

direct measurement approaches. In summary, for a given system, it appears direct measurement of solubility may require a high-level optimisation of experimental methodology to ensure reliable estimates under the conditions of interest.

Aside from potential practical challenges, like any experimental approach, the effort (in terms of time, skill and economic cost), as well as the amount of material required to perform a robust experimental design is a common critique of experimental solubility approaches. However, attempts have been made to design methodologies where common solubility measurement techniques are completed automatically, leading to high-throughput approaches, when compared to human-operated experiments [25, 26, 27]. Although practical challenges with solubility measurements have been discussed in the literature since 1952 [18], and described in varying levels of detail in widely available textbooks [6, 19], it is interesting to note that issues are still common in published work [15, 20].

Model-based Approaches

Challenges with purely experimental approaches for the estimation of solubility motivate the use of model-based approaches. One model-based approach involves correlating directly measured solubility with a mathematical model, after which they can be used to interpolate between conditions where solubility was measured and, in some cases, extrapolate to conditions where solubility is yet to be determined. This approach is particularly attractive for cases where reliable directly measured solubility data is available across the entire range of interest. Within this approach, the rationale for models used varies [6]. For example, in some cases models derived from a mathematical perspective (e.g., polynomial expansions) are used which have no connection to an underlying theory of solubility beyond the assumption that, in a certain range, it can be expressed as a continuous function. In other cases, models are derived from physical theories (c.f. Section 2.2).

The correlative approach has been discussed extensively in [28], where fifteen analytical equations with two, three or four flexible parameters describing solubility as a function of temperature were fit to solubility data across a variety of temperature intervals for forty one organic compounds in a broad range of organic solvents and water. As an aside, the collection of so-called 'regression equation(s) of solubility' (RES) included examples of models with thermodynamic basis, as well as arbitrary mathematical forms In terms of correlating temperature dependent solubility data, generally speaking, models with four parameters gave the best goodness of fit, taken here as the χ^2 summed over all models in a specific class, with the exception of organic solutes in water exhibiting negative deiations from Raoult's Law at saturation, where models with three parameters gave comparable or sometimes marginally better performance. This is generally expected; however, could indicate over-fitting. The ability for each model to reliably extrapolate beyond the range of temperature over which solubility was measured was assessed by predicting the pure crystal

solute melting temperature by fixing the mole fraction solubility to one. It was found that the best performing class models were two and three parameter equations, giving melting temperature predictions with average absolute errors of 19.8 and 21.7 K respectively. Interestingly, four parameter models predicted a positive melting temperature in only 61.8, 55.9, 64.7 and 55.9 % of cases (for each four parameter model, respectively) which further suggests over-fitting. In summary, analysis indicates that, when possible, collecting solubility data across a broad range of temperature and correlating with an appropriate model can lead to reasonable results in terms of interpolation and extrapolation. However, regardless of the chosen model(s), correlating directly measured solubility data requires reliable solubility data across a broad range of temperature which, as discussed already, is prone to several context specific challenges. Furthermore, based on typical experimental designs where the solubility is reported over a narrow range of temperature around ambient conditions (e.g., 280–330 K), it is unclear how reliable a given extrapolation will predict solubility for a given experimental design, model and system. This uncertainty can, of course, be reduced if the solubility was measured at higher temperatures; however, this has potential challenges as discussed already.

Another model-based approach for the estimation of solubility are those based on Quantitative Structure Property Relationships (QSPR) [29, 30]. Here, models are developed by correlating large databases of solubility data to so-called 'molecular descriptors', from which it is then possible to predict the solubility of compounds not included in the initial model development. In comparison to a correlative application, these approaches are attractive given that they can, in principle, be applied to many systems, even those not included in the training, which means no additional solubility data is required, in principle. Quantitative structure property approaches accurately predict relative solubility of a single solute in several solvents or multiple solutes in the same solvent accurately, as such their application is likely more suitable for screening purposes, either during early stage drug design or solvent selection in the pharmaceutical industry. As reviewed in [31], examples of the QSPR approach applied to solubility are the so-called 'General Solubility Equation (GSE)' [32] and the Linear Solvation Energy Relationship (LSER) [33], and a recent example where machine learning approaches were used in a QSPR framework can be found in [34].

For the purpose of this thesis, a thermodynamics based approach to estimate solubility is considered to be one where the model used is based in general framework of equilibrium thermodynamics. An attractive property of thermodynamics-based approaches to the estimation of solubility is that thermodynamics provides a theoretical framework to estimate solubility with limited, and in some cases no, solubility data. It should be noted that analytical forms derived from thermodynamic arguments are sometimes used for correlative purposes [6]; however, parameters that have a thermodynamic interpretation are usually treated as flexible fitting parameters, obscuring their meaning. Thermodynamics-based approaches for the estimation of solubility are discussed

in greater detail in Section 2.3, following a description of the fundamentals of thermodynamics in the next section.

2.2 Fundamentals of Equilibrium Thermodynamics

Thermodynamics or, more specifically, classical equilibrium thermodynamics, is a framework originally developed to describe empirically observed relationships between heat and work at macroscopic length and time scales, specifically in the context of heat engines [35]. Over time, thermodynamics was extended to an all-encompassing physical theory of energy and matter, described by the Laws of Thermodynamics (i.e., Zeroth, First, Second and Third), which are widely accepted in the absence of contrary experience. A key result from classical equilibrium thermodynamics is a description of phase equilibrium, which allows the theoretical boundaries (i.e., solid-liquid equilibrium curves) governing crystallisation processes to be expressed as mathematical equations.

The mathematical framework of thermodynamics is concisely presented through four postulates stated in [5], from which all properties of thermodynamics relevant to this thesis are derived. The postulates are given as:

Postulate I There exists particular states (called equilibrium states) that, macroscopically, are characterised completely by the specification of the internal energy (U) and a set of extensive parameters X_1, X_2, ..., X_t, specifically numerated for a given type of system.

Postulate II There exists a function called entropy (S) of the extensive parameters, defined for all equilibrium states, and having the following property. The values assumed by the extensive parameters in the absence of a constraint are those that maximise entropy over the manifold of constrained equilibrium states.

Postulate III The entropy of a composite system is additive over the constituent subsystems (whence the entropy of each constituent system is a homogeneous first-order function of the extensive parameters). The entropy is continuous and differentiable and is a monotonically increasing function of the energy.

Postulate IV The entropy of any system vanishes in the state for which
$T \equiv (\partial U/\partial S)_{X_1,X_2,\ldots} = 0$.

As an aside, an extensive property is defined as one which scales with the 'size' of the system, in contrast to an intensive property which is independent of 'size'. Beyond the postulates, essential to the development of thermodynamics is the generalisation of the principle of energy conservation to include heat and work giving the First Law of thermodynamics stated mathematically as:

$$dU = \delta Q - \delta W, \tag{2.2}$$

where Q and W denote heat and work (all types), respectively. Restricting analysis to systems where the relevant extensive parameters include: internal energy (U), volume (V) and mole number of chemically distinct species (N_1, N_2, ..., N_n), the postulates give:

$$U = U\left(S, V, N_1, N_2, \ldots, N_n\right), \tag{2.3}$$

which becomes a so-called 'fundamental property relation' [5, 35, 36] describing an infinitesimal thermodynamic change:

$$dU \leq TdS - PdV + \sum_{i}^{n} \mu_i dn_i, \tag{2.4}$$

where T is temperature, P is pressure and μ_i is the chemical potential, all of which are intensive properties. Comparison of Eq. (2.2) and (2.4) gives the Second Law of thermodynamics, which defines changes in entropy as $dS \geq (\delta Q/T)_{V,n_j}$. The term PdV comes from the specification that the only type of work is that due to expansion. From Eq. (2.4), it should be noted that given a function $U = U(S, V, N_1, \ldots, N_n)$ for a specific system, all thermodynamic properties of can be derived.

As an aside, it can be shown [5]:

$$\sum_{i}^{N} n_i d\mu_i = VdP - SdT, \tag{2.5}$$

which, at fixed temperature and pressure, relates the chemical potential of components in a phase. This relation is referred to as the Gibbs-Duhem equation and can be used to check consistency of of experimental data [36]. For a binary system (i.e., two components), if the chemical potential of one component is known it is possible to estimate the chemical potential of the other up to a constant of integration, given that:

$$d\mu_A = -\frac{n_B}{n_A} d\mu_B. \tag{2.6}$$

Typically, thermodynamic systems comprise multiple phases, each of which may comprise more than one component. The postulates alone define how intensive properties across a system behave at equilibrium. For an isolated system (i.e., unable to exchange energy and matter with it's surroundings) comprising two open systems (i.e., free to exchange energy and matter), the change in entropy is given by:

$$dS = \frac{1}{T^\alpha} dU^\alpha - \sum_{i}^{n} \frac{\mu_i^\alpha}{T^\alpha} dN_i^\alpha + \frac{P^\alpha}{T^\alpha} dV^\alpha + \frac{P^\beta}{T^\beta} dV^\beta + \frac{1}{T^\beta} dU^\beta - \sum_{i}^{n} \frac{\mu_i^\beta}{T^\beta} dN^\beta, \tag{2.7}$$

but noting that, because the composite system is isolated, $dU^\alpha = -dU^\beta$ and $dN_1^\alpha = -dN_1^\beta$ and

$dV^\alpha = -dV^\beta$, while $dS = 0$ (from Postulate 2) Eq. (2.7) gives:

$$T^\alpha = T^\beta \tag{2.8}$$

$$P^\alpha = P^\beta \tag{2.9}$$

$$\mu_1^\alpha = \mu_1^\beta \tag{2.10}$$

$$\dots \tag{2.11}$$

$$\mu_n^\alpha = \mu_n^\beta \tag{2.12}$$

which states that temperature, pressure and chemical potential (of each individual component in each phase) are equal at thermodynamic equilibrium. It should be noted that postulate II leads directly to the statement that internal energy is minimised when a system is at thermodynamic equilibrium (c.f. [5]).

As an aside, the 'Gibbs phase rule' defines the degrees of freedom (i.e., independent state variables) required to specify the state of a non-reacting multi-phase system:

$$F = 2 + N - \pi, \tag{2.13}$$

where N is the number of chemical species present in the system and π is the number of phases.

Understanding the behavior of multi-phase, multi-component systems is crucial in chemical engineering. For example, consider distillation—where the phase composition of a multi-component vapor-liquid system is changed by sequentially varying it's temperature and pressure. Knowing how the composition of both vapor and liquid phases at equilibrium allows us to design, develop and optimize distillation processes. From the postulates and the principle of energy conservation, a vast number of chemical engineering problems can be solved, or at least described, if the function $U = U(S, V, N_1, \dots)$ for the system of interest was known. However, entropy and volume are generally not easily controllable process parameters—temperature and pressure are more meaningful. In the next section, this problem is resolved by the introduction of another potential defined as the Gibbs free energy, which plays a significant role in every chapter of this thesis. Some relevant properties are also explored.

2.2.1 The Gibbs Free Energy

Thermodynamic potentials are defined (mathematically) by taking Legendre transforms of Eq. (2.4). The Gibbs free energy (G) is an example of a thermodynamic potential, defined in differential form for an open system capable of expansion work only:

$$dG \leq -SdT + VdP + \sum_i^n \mu_i dn_i. \tag{2.14}$$

The Gibbs free energy is expressed in terms of intensive properties T and P both of which are, crucially, easier to control compared to entropy when performing experiments. Like internal energy, the

Gibbs free energy of a system is minimised at equilibrium. In addition, if $G = G(T, P, N_1, \ldots, N_n)$ is known for a given system, then it is completely specified and it is possible to estimate any thermodynamic quantity.

The extensive Gibbs free energy of single phase, multi-component thermodynamic system can be expressed as as a mole number weighted sum of the chemical potential:

$$G^{\alpha} = \sum_{i}^{n} n_i^{\alpha} \mu_i^{\alpha}, \tag{2.15}$$

which, gives the Gibbs free energy of a multi-phase, multi-component system as:

$$G^{\text{system}} = \sum_{\alpha}^{p} \sum_{i}^{n} n_i^{\alpha} \mu_i^{\alpha}, \tag{2.16}$$

where p is the number of phases.

At fixed pressure and composition, changes in the Gibbs free energy with temperature can be related to thermal properties. At constant temperature and pressure, the Gibbs free energy is related to enthalpy through the Gibbs-Helmholtz relation:

$$\left(\frac{\partial \beta G}{\partial \beta}\right)_{P, n_T} = H, \tag{2.17}$$

where $\beta = (RT)^{-1}$ and introducing enthalpy (another thermodynamic potential), defined for an infinitesmal change as:

$$dH \leq T dS + V dP + \sum_{i}^{N} \mu_i dn_i. \tag{2.18}$$

Going further, the Gibbs free energy is related to constant pressure heat capacity (i.e., $(\partial H/\partial T)_P \equiv C_p$):

$$\left(\frac{\partial^2 \beta G}{\partial T \partial \beta}\right)_{P, n_T} = C_p. \tag{2.19}$$

At fixed temperature and composition, changes in the Gibbs free energy with pressure can be related to volumetric properties:

$$\left(\frac{\partial G}{\partial P}\right)_{T, n_T} = V. \tag{2.20}$$

From Eq. (2.15), the extensive Gibbs free energy of a homogeneous phase at fixed temperature and pressure can be expressed as a mole number weighted sum of the chemical potential of all species. However, as already stated, the chemical potential is not directly measurable. It is customary to interpret the chemical potential with reference to an auxiliary quantity defined as fugacity which, at constant temperature, defines the chemical potential as:

$$\mu_i = \mu_i^{\text{ref}} + RT \ln \frac{f_i}{f_i^{\text{ref}}}, \tag{2.21}$$

where μ_i^{ref} and f_i^{ref} are the chemical potential and fugacity of component i in a chosen reference state and f_i is the fugacity of component i in the 'real' state. Crucially, the introduction of fugacity reduces Eq. (2.12) to:

$$f_i^{\alpha} = f_i^{\beta}. \tag{2.22}$$

Combining Eq. (2.15) and (2.21), for a generic multi-component phase, gives:

$$G = \sum_i n_i \left(\mu_i^{\text{ref}} + RT \ln \frac{f_i}{f_i^{\text{ref}}} \right) \tag{2.23}$$

$$= G^{\text{ref.}} + G^{\text{cor.}} \tag{2.24}$$

where $G^{\text{ref.}}$ is a contribution to the extensive Gibbs free energy associated to the reference states chosen for all species and $G^{\text{cor.}}$ is a contribution giving a correction to the reference state. Different reference states have been proposed for different states of matter based on physically observed limiting behavior of real systems (e.g., ideal gas and ideal solution (Raoult and Henry)). As an aside, it is straightforward to derive the Gibbs free energy of an ideal gas as well as an ideal solution (defined with respect to Raoult's Law) from statistical mechanics by assuming they are comprised of either non-interacting (in the case on an ideal gas) or identically interacting (in the case of a solution (i.e., cross interactions are the same as like interactions)) particles, yielding the same empirically observed relationships.

For liquid mixtures, it is customary to interpret the Gibbs free energy using Eq. (2.24). Although many references states exist, perhaps the most widely applied is that of the 'ideal solution' defined with respect to Raoult's Law (c.f. Henry's Law). Briefly, Raoult's Law comes from empirical observations indicating that under certain conditions the pressure of a vapor in equilibrium with a liquid is given by:

$$P(T) = x_1 p_1^{\text{sat.}}(T) + x_2 p_2^{\text{sat.}}(T) + \ldots, \tag{2.25}$$

where x_i is the mole fraction of component i in the liquid phase and $p_i^{\text{sat}}(T)$ is the vapor pressure of component i at temperature T. This leads to the definition of the Gibbs free energy of an ideal solution given by:

$$G^{\text{id}} = \sum_i^N n_i \mu_i^{l,\circ}(T, P) + \sum_i^N n_i RT \ln x_i \tag{2.26}$$

where the chemical potential of a component in an ideal solution is given by $\mu_i = \mu_i^{l,\circ} + RT \ln x_i$, with $\mu_i^{l,\circ}$ denoting the chemical potential of component i in the pure liquid state at T and P.

If $G^{\text{ref.}}$ in Eq. (2.24) is taken as the Gibbs free energy of an ideal solution at the same temperature and pressure as the real phase, $G^{\text{corr.}} = G^{\text{E}}$ where G^{E} is the 'excess' Gibbs free energy. For real solutions, the chemical potential is defined by introducing the activity coefficient γ_i, giving:

$$\mu_i(T, P, x) = \mu_i^{l,\circ}(T, P) + RT \ln \left(x_i \gamma_i(T, P, x) \right), \tag{2.27}$$

which gives another definition for the excess Gibbs free energy as:

$$\beta G^E = \sum_i^N n_i \ln \gamma_i. \tag{2.28}$$

2.2.2 Thermodynamic Modelling

As will be illustrated in the proceeding chapters, estimates of the underlying Gibbs free energy of different phases (i.e., pure crystal, solute-solvent liquid mixtures, etc.), are required to make solubility predictions via thermodynamics given that, as described previously, if the Gibbs free energy for a given system as a function of temperature, pressure and composition, then all other thermodynamic properties of said system can be predicted.

It is not possible to measure the Gibbs free energy directly; however, it is possible to relate the underlying Gibbs free energy to thermodynamic properties that can be measured. As an example, consider the Gibbs free energy of a phase a fixed composition. Through Eq. (2.17)–(2.20), it is possible to relate changes in the Gibbs free energy (relative to an reference state) to experimentally measurable thermal and volumetric properties (i.e., heat capacity and density) through:

$$\beta G = \beta_0 G_0 + \beta_0 H_0 \left(\frac{\beta}{\beta_0} - 1 \right) + \int_{\beta_0}^{\beta} \int_{T_0}^{T} C_p(P_0) dT d\beta + \int_{P_0}^{P} V(T) dP, \tag{2.29}$$

where G_0 and H_0 are the extensive Gibbs free energy and enthalpy defined at the initial conditions (i.e., T_0 and same pressure and composition) and where the 'state function' properties of the Gibbs free energy was used. An expression analogous to Eq. (2.29) is available for changes in the chemical potential at fixed composition:

$$\beta \mu_i = \beta_0 \mu_{i,0} + \beta_0 \bar{h}_{i,0} \left(\frac{\beta}{\beta_0} - 1 \right) + \int_{\beta_0}^{\beta} \int_{T_0}^{T} \bar{c}_{p,i}(P_0) dT d\beta + \int_{P_0}^{P} \bar{v}_i(T) dP, \tag{2.30}$$

where $\bar{h}_i$, $\bar{c}_{p,i}$ and $\bar{v}_i$ are the the partial molar enthalpy, constant pressure heat capacity and volume of component i, respectively.

Equation (2.29) and (2.30) relate changes in abstract quantities for which no experimental measurement techniques exist to measurable thermal and volumetric quantities, defined as partial derivatives of the Gibbs free energy [36]. The integrals can be evaluated numerically from tabulated experimental measurements or analytically if closed form expressions for $C_p(T, P, n)$ and $V(T, P, n)$ are known, in addition to $G_0(T, P, n)$ and $H_0(T, P, n)$.

Constructing a mathematical model for a phase with fixed composition is straightforward if the heat capacity and density are known as a function of temperature and pressure, respectively. However, modeling phases that vary in composition is more challenging, since the composition partial derivatives of the Gibbs free energy are the chemical potential (and subsequent higher order partial derivatives), none of which are directly measurable. Instead, all available experimental data for the system under study is collected and correlated using different mathematical models until an appropriate model is found that describes available experimental data deemed reliable enough to do so. Liquid mixtures comprising a solute and solvent are the primary composition dependent phase encountered throughout the remainder of this thesis and so the subsequent discussion is focused on the development of models for this type of phase specifically.

As suggested already, it is common to interpret liquid mixtures in terms of an ideal solution contribution (defined with respect to Raoult's Law) and an excess Gibbs free energy contribution (c.f. equations of state [35]). In this framework, the ideal contribution is modelled by pure liquid component properties in a manner similar to Eq. (2.29), leaving the excess Gibbs free energy. There has been significant research effort towards development of models for the excess Gibbs free energy and interested readers should refer to seminal texts, such as [36], for a detailed account describing the historical development of several. Regardless, examples describing non-electolyte liquid mixtures include: the Van Laar [37], Scatchard-Hildebrand [38, 39], Redlich-Kister [40], Wilson [41], NRTL (**N**on-**R**andom **T**wo **L**iquid) model [42] and UNIQUAC (**UNI**versal **QUA**si-**C**hemical) [43] models. It should be noted that references to works where the excess Gibbs free energy model was originally proposed are provided; however, models are routinely modified to improve their ability to simultaneously describe multiple types of experimental data, as well as their ability to extrapolate to regions where data is not available.

Typically, the most appropriate model structure, as well as the corresponding value of empirical parameters in the aforementioned excess Gibbs free energy models are developed by correlating available experimental data for the system(s) under study. Briefly, correlating a model to experimental data involves finding model structures and corresponding sets of parameter values that 'best' describe the available data. This general approach is applied throughout the remainder of this thesis. The modelling approach and computational details vary across chapters, as such, details of methods used are provided where relevant.

Given that, generally speaking, appropriate coverage, type, and accuracy of experimental thermodynamic data such that it is possible to construct reliable Gibbs free energy models for phases over the entire range of temperature, pressure and composition is rarely available for a given system, considerable effort has been gone towards the development of so-called 'group contribution' approaches that attempt to by-pass this issue. Here, the properties of a molecule are estimated as a combination of the properties of the functional groups that comprise the molecule's structure. Once a list of functional groups has been determined, each functional group is assigned an unknown contribution to the property, which is then evaluated by correlating experimental data for the properties of interest for a broad range of compounds including as many functional groups as possible. Once the contribution of each functional group is known, it is possible to re-build molecules that weren't included in functional group parameter estimation and prediction the desired property, with the goal being the reduction in experimental effort required to analyse new systems. The group contribution approach has been applied to many thermodynamic properties [44]; however, perhaps most relevant to the determination of the excess Gibbs free energy (and activity coefficients) [45] through the UNIFAC (**UNIQUAC F**unctional-group **A**ctivity **C**oefficients) method [46] (and modifications [47]), which re-interpretted the aforementioned UNIQUAC excess Gibbs free

energy model. The development of functional group parameter databases for UNIFAC and Modified UNIFAC group contribution models is primarily completed on a large scale by the UNIFAC consortium using data collated in the Dortmund Data Bank. However, parameter databases are also developed on a much smaller scale in independently published works, where authors have proposed new functional groups, parameter values and parameter dependencies, to make the general framework more reliable for a sub-set of multi-component systems (e.g., Pharma UNIFAC [48] for pharmaceutical systems and S-UNIFAC [49] for sugar solutions).

In addition to group contribution methods, there are models based on the so-called 'conceptual segment contribution concept' [50], originally applied to the aforementioned NRTL model. Rather than interpreting a molecule in terms of functional groups, the conceptual segment contribution concept views a molecule in terms of conceptual segments related to interactions between molecules in solution, including: hydrophobic, polar (repulsive and attractive) and hydrophilic. Once molecules have been interpreted in terms of their respective segments, a characteristic value for each segment contribution is determined by correlating experimental data and then predictions can be made for molecules not included in the initial parameterisation. The conceptual segment contribution concept has been applied to the UNIQUAC model, giving the UNIQUAC-SAC model [51].

More recently, as noted in [52], the general method leveraged in group contribution, conceptual segment contribution and QSPR-based approaches where molecular descriptors are correlated to physical properties has been approached using Machine Learning (ML) techniques. See also a recent review in [53].

Generally speaking, the predictive performance of group contribution and conceptual segment contribution models is highly system dependent, as well as the database used to generate model parameters.

In the approaches described thus far, the identification of appropriate model structure, as well as parameter values, is done by correlating experimental data (regardless of whether parameters can be transferred to new systems as done in group-contribution and conceptual segment-based frameworks). Beyond bespoke parameterisations as well as group contribution (and related) methods, it should be noted that progress has been made towards *a priori* predictive methods, specifically for the estimation of the excess Gibbs free energy of components in liquid mixtures through the COSMO-RS [54] and COSMO-SAC [55] methods. Here, all that is required to make predictions for thermodynamic properties of liquid mixtures is the molecular structure of the compounds in the mixture, as well a set of pre-determined, compound-independent parameters [56]. Furthermore, it is possible to predict model parameters *in silico* using Molecular Dynamics (MD) simulations, as shown in [57], where binary interaction parameters in the UNIQUAC model were predicted for sugar-water systems.

2.3 Estimating Solubility in the Context of Thermodynamics: A Review

It is convenient to begin the description of solubility in the context of equilibrium thermodynamics by considering a closed system comprising two phases; one consisting of a pure crystalline substance (i.e., solute) and another consisting of a liquid comprising dissolved solute and solvent (i.e., solution). Assuming the composite system is at thermodynamic equilibrium, by definition, the solution is saturated. As described already, a multi-phase system exists at equilibrium when the Gibbs free energy is minimized, which also leads to the following condition:

$$\mu_i^{\mathrm{s},\circ}(T) = \mu_i^{\mathrm{l,sat}}\left(T, x_{\mathrm{i,sat}}(T)\right), \tag{2.31}$$

where, as before, $\mu_i^{s,\circ}$ and $\mu_i^{\mathrm{l,sat}}$ are the chemical potential of solute in the pure crystalline phase and saturated liquid solution, respectively. Equation (2.31), is the starting point for perhaps the most widely applied predictive model referred to in crystallisation literature and will be discussed further in the following section.

2.3.1 Analytical Models Based on the Normal Melting Point

If, for example, the chemical potential of solute in liquid solution is interpreted relative to Raoult's Law, Eq. (2.31) becomes:

$$\ln x_i\gamma_i = \beta\left(\mu_i^{\mathrm{s},\circ} - \mu_i^{\mathrm{l},\circ}\right), \tag{2.32}$$

where x_i is the mole fraction of solute in solution, γ_i is the solute activity coefficient in solution, $\mu_i^{\mathrm{l},\circ}$ is the chemical potential of the solute as a pure liquid, and $\beta = 1/RT$. It should be noted that the pure liquid solute is usually referred to as a 'hypothetical state', given that the solute's thermodynamically stable state is the pure crystal at all temperatures along the soubility curve.

Focusing specifically on $\mu_i^{\circ,\mathrm{s}} - \mu_i^{\circ,\mathrm{l}} = \Delta\mu_i$ which doesn't depend on composition (by definition), it can be shown that:

$$\beta\Delta\mu_i(T) = \beta_0\Delta\mu_i(T_0) + \Delta h_i(T_0)\left(\frac{\beta}{\beta_0} - 1\right) + \int_{\beta_0}^{\beta}\int_{T_0}^{T}\Delta c_p\, dT\, d\beta \tag{2.33}$$

where Eq. (2.17) and (2.19) have been used. When the reference temperature is taken as the assumed pure solute melting temperature, Equation (2.33) reduces to:

$$\beta\Delta\mu_i(T) = \Delta h_i(T_m)\left(\frac{T_m}{T} - 1\right) + \int_{\beta_m}^{\beta}\int_{T_m}^{T}\Delta c_p\, dT\, d\beta \tag{2.34}$$

noting that $\beta\left(\mu_i^{s,\circ}(T_m) - \mu_i^{l,\circ}(T_m)\right) = 0$. Finally, combining Eq. (2.32) and (2.34), an explicit equation for solubility is found:

$$\ln x_i = \Delta h_i(T_m)\left(\frac{T_m}{T} - 1\right) + \int_{\beta_m}^{\beta}\int_{T_m}^{T}\Delta c_p\, dT\, d\beta - \ln\gamma_i, \tag{2.35}$$

where $\Delta h_i(T_m)$ is the molar enthalpy associated to the melting of pure crystal solute and $\Delta c_p = c_p^l - c_p^s$ where c_p^s and c_p^l are the molar heat capacity of the solute in as a pure solid and liquid, respectively.

Equation (2.35) represents an exact[1] model describing the temperature dependence of solubility. By solving Eq. (2.35), it is possible to estimate solubility anywhere if $\Delta h_i(T_m)$, $\Delta c_p(T)$ and $\ln \gamma_i(T, x)$ are known, in principle. Given that there exist experimental measurement and predictive techniques for the estimation of pure solute melting properties (i.e., enthalpy and melting temperature) and heat capacity, and approaches to estimate the solute activity coefficient, Eq (2.35) has appealing properties.

Equation (2.35) is rarely, if ever, used in practice, and a variety of simplifications are applied yielding approximate equations discussed in various contemporary texts that treat solubility in the context of thermodynamics [3, 6, 36]. Historically, approximations were in part motivated by the lack of direct approaches to estimate Δc_p [58], and more specifically, the heat capacity of the hypothetical pure liquid solute at temperatures below the melting point. Some common approximations include a case where Δc_p is assumed to be a constant, independent of temperature, giving:

$$\ln x_i \approx \Delta h_i(T_m) \left(\frac{T_m}{T} - 1 \right) + \frac{\Delta c_p}{R} \left(\ln \frac{T_0}{T} - \frac{T_0}{T} + 1 \right) - \ln \gamma_i, \tag{2.36}$$

while, if Δc_p is assumed to well approximated by the entropy of fusion first proposed by Hildebrand, Eq. (2.35) becomes:

$$\ln x_i \approx -\frac{\Delta h_i(T_m)}{R T_m} \ln \frac{T_m}{T} - \ln \gamma_i, \tag{2.37}$$

while, if Δc_p is ignored completely, Eq. (2.35) becomes:

$$\ln x_i \approx \Delta h_i(T_m) \left(\frac{T_m}{T} - 1 \right) - \ln \gamma_i. \tag{2.38}$$

As an aside, the contribution to Eq. (2.35) describing the temperature dependence of the solute melting and associated approximations can be interpreted as a Taylor series expansion (c.f. Chapter 3). Thus, for systems with solutes whose melting point is close to ambient conditions as well as close to ideal behavior in the solution phase, it might be expected that the approximations are valid from a mathematical perspective.

If the activity coefficient in Eq. (2.35) is ignored which can be interpreted as the solution behaving ideally (relative to Raoult's Law), then solubility predictions made have been referred to as the ideal solubility. The validity of approximations has been discussed in the context of estimating the ideal solubility [59, 60, 61] and summarised in [62]. However, the impact of approximations is perhaps best illustrated in [58], where it has been shown that assumptions regarding the behavior of Δc_p on estimated ideal solubility at 298.15 K can have a significant impact for compounds with melting points in the range 420–450 K. The authors measured the melting

[1] Ignoring physical effects due to surface forces, as well as electric, magnetic and gravitational fields [36].

enthalpy and temperature of five compounds (i.e., acetaminophen, naproxen, anisic acid, mannitol and diethylstilbestrol), and compared predicted ideal solubility at 298.15 K using Eq. (2.36), 2.37 and (2.38), ignoring the term associated to the solute activity coefficient in solution. The authors noted that differences in approximation led to differences in ideal solubility at 298.15 K ranging from 59–189% (mole fraction solubility). In summary, it is unclear how, for a given system, the choice of approximation might impact reliability of solubility predictions.

Although approximations are usually employed to handle the contribution associated to the difference in heat capacity between pure crystal and 'hypothetical' liquid solute, a recent work [63] has reported the heat heat capacity of pharmaceutical compounds, including: (RRS)-ibuprofen, (S)-ibuprofen, (S)-naproxen and indomethacin, in both pure crystal and sub-cooled liquid states at temperatures above and below the normal melting point, potentially indicating that reliable estimation if this term is possible. In recent work [64], it was shown that assuming the heat capacity to be a constant, independent of temperature, is a reasonable assumption for certain compounds, illustrated by interpolation of heat capacity for solutes in chemically and physically similar solvents in combination with enthalpy of fusion and solution measurements.

For predictive purposes (i.e., avoiding direct solubility measurement), the enthalpy of fusion is required, which is frequently reported in experimental studies. However, as described in [65], measuring the enthalpy of fusion is limited for compounds with poorly defined melting properties, which can be due to thermal decomposition or solid form transition prior to melting, an example being the simple amino acid glycine, which has been shown to decompose through a combination calorimetry, thermogravimetry and mass spectroscopy. Regardless, it should be note that that novel approaches have recently been proposed based on 'Fast Scanning Calorimetry' (FSC) [66] which, like the hypothetical pure liquid solute, potentially allow determination of thermodynamic quantities previously thought to be inaccessible. More generally, even when a compound appears to have a well defined melting transition, accurate determination of melting properties can be made challenging due to "noncrystalline (amorphous) residues, large molecular sizes, highly variable crystal sizes, defects in crystal structures, impurities, superheating, isomerization and mutarotation, and [...] the presence of water (solvent) at normal atmospheric surroundings" [67], which are known to have a meaningful impact on measured melting property values.

For reliable prediction of solubility, application of Eq. (2.35) requires knowledge of the solute activity coefficient as a function of temperature and composition and, more specifically, along the solubility curve of a given system. As described already in Section 2.2.1, there exists an enormous body of literature on the estimation of activity coefficients, given that they are required in any practical application involving multi-component liquid mixtures. For the prediction of solubility via Eq. (2.35), predictions are often viewed as a two-step process, where pure solute melting parameters (i.e., $\Delta h^{\text{fus}}(T_m)$ and T_m) are evaluated experimentally, $\Delta c_p(T)$ is approximated in any

one of the of many ways described already (or ignored), and the solute activity coefficient is estimated using an excess Gibbs free energy model parameterised in 'outside' of the solubility prediction—either using previously published parameters or a bespoke parameterisation based on available experimental data.

For example, in [68], the original UNIFAC group-contribution method for the estimation of activity coefficients in multi-component liquid mixtures was used to predict the solubility of various non-electrolyte solids in single and mixed solvents was predicted using Eq. (2.38) where the contribution associated to the activity coefficient of solute in solution was modelled using the UNIFAC group-contribution method. For the systems investigated, predicted solubility was found to be in good agreement with experimental data from the melting point to the eutectic point; however, it should be noted that solids investigated had a normal melting temperature below 373.15 K, which is, relatively speaking, low when compared to industrially relevant compounds.

In a more recent study [69], the original UNIFAC group-contribution method was again used to estimate solute activity coefficients for solubility prediction via Eq. (2.38). Compared to [68], the systems investigated included compounds with melting temperatures in excess of 373.15 K, although the selected solvents were similar. It was shown that predictions at 298.15 K were in poor agreement with experimental data, with the error in predicted solubility in excess of 15% for most systems and deemed too unreliable for the design of industrial crystallisation processes. The authors also highlight that differences in solubility predictions based in i) ignoring the Δc_p term and ii) assuming it to be a constant, can lead to substantial (e.g., up to 100%) differences in the solubility predicted at 298.15 K.

In [70], several excess free energy models (i.e., UNIFAC, Modified UNIFAC, COSMO-SAC and NRTL-SAC) were used to predict the temperature dependent solubility of ibuprofen, acetaminophen, benzoic acid, 4-aminobenzoic acid and anthracene in a variety of solvents using Eq. (2.38) and (2.37). The authors concluded that, even when factoring in the Δc_p term, predictions made were generally not suitable for quantitative prediction required for the design of crystallisation processes. It was suggested that this is due to the excess Gibbs free energy models and associated parameter values used.

In a very recent work (i.e., 2022) [56], reviewed solubility for predictions for ten pharmaceutical compounds (e.g., paracetamol, (RS)-ibuprofen, carbamazepine, etc.,) in a variety of common industrial solvents using Eq. (2.35) assuming $\Delta c_p(T) = \Delta c_p(T_m)$ and the solute activity coefficient predicted using both PC-SAFT and COSMO-RS modelling frameworks. Compared to studies mentioned previously, the solid compounds investigated are higher in molecular weight as well as melting temperature. Although PC-SAFT has semi-empirical parameters that are usually evaluated by correlating experimental data, the author used it in a 'predictive manner' using pure component properties, only. As described, solubility predictions were generally in qualitative agreement with

direct measurements for the systems studied and both models with predictions based on solute activity coefficients estimated using COSMO-RS found to be in better agreement. However, predictions based on both models were in poor quantitative agreement with direct measurements, in general agreement with findings for predictions based on other models describe already.

It should be noted that Eq. (2.35) (in approximate form) is used to parameterise excess Gibbs free energy models if temperature dependent solubility is available across a broad range of temperature, as well as the melting properties of the system under study. For example, in [71], the functional group interaction parameter values for the SAFT-γ Mie equation of state were evaluated by correlating temperature dependent solubility data for benzoic acid, diphenylamine and mefenamic acid in a range of solvents.

2.4 Conclusions

In this chapter, the fundamentals of crystallisation were introduced and the importance of solubility as an essential property for the rational design, development and optimisation of crystallisation processes highlighted. As described, it is possible to measure solubility directly using any one of many technique-instrument combinations available, in principle. However, although apparently straightforward, direct determination of solubility can be challenging, as well as costly in terms of time and resources.

Model-based approaches for the estimation of solubility were introduced, including correlative, QSPR, as well as those based in equilibrium thermodynamics. The relevance of equilibrium thermodynamics as a framework to model phase equilibria and, more specifically, solubility, was elucidated by a considering the fundamentals of equilibrium thermodynamics in the postulative framework formulated in [5].

The most common approach found in crystallisation literature where an explicit model for solubility expressed in terms of pure solute melting properties, differential heat capacity and solution-based activity, is derived from fundamental thermodynamic relationships and it's application reviewed. As described, this model is usually applied with simplifying assumptions, the impact of which are generally not well known on system-by-system basis. In addition, the majority of applications available in literature approximate the solute activity coefficient using 'predictive' excess Gibbs free energy models (e.g., UNIFAC, NRTL-SAC, COSMO-RS), the reliability of which is also challenging to assess for a given system. Finally, this approach relies on knowledge of the pure solute enthalpy of fusion and melting temperature, and solution non-ideality, all of which can be challenging for a variety of reasons discussed.

From the above, the following is noted. First, there is scope to develop methods that allow estimation of solubility as a function of temperature that do not require measurement of pure solute

melting properties (i.e., $\Delta h^{\text{fus.}}(T_m)$ and T_m) that provide alternative approaches for systems where this is challenging. Second, there is scope to develop methods that allow estimation of solubility as a function of temperature without the need to employ simplifications, the implications of which are hard to assess on a given system in the absence of direct solubility data, whether that be through the introduction of simplifying assumptions, or reliance on 'predictive' methods for excess Gibbs free energy properties. These areas are the subject of work presented in the following chapters.

Chapter 3

A Local Expansion Approximation of the van't Hoff Equation

The results presented in this chapter have been published in the journal Crystal Growth and Design (doi: 10.1021/acs.cgd.1c01217), authored by Andrew Manson, Jan Sefcik and Leo Lue. It should be noted that large sections of this chapter reflect those in the published manuscript; however, adaptations have been made to improve readability in the context of this thesis.

3.1 Introduction

In Chapter 2, it was shown that an explicit model describing the temperature dependence of solubility for a given solute in solution could be derived from fundamental thermodynamics by referring to the pure solute normal melting point. In this chapter, we construct an approximate model describing solubility as a function of temperature by considering the so-called 'van't Hoff' equation, which describes the temperature dependence of thermodynamic processes. More specifically, we approximate the temperature and composition dependence of thermodynamic quantities in the van't Hoff equation applied to dissolution using 1^{st}-order Taylor series expansions. As shown, this gives an approximate model from which solubility can be predicted if relevant thermodynamic quantities are known, all of which can be approximated from experimentally measurable properties and, crucially, none of which include the pure solute melting properties.

Using this approach, we investigate the aqueous solubility of α, β and γ-glycine which, despite numerous studies, have considerable quantitative uncertainty, in particular for the most stable (γ) and the least stable (β) solid forms. The remainder of this chapter is structured as follows. In the next section, the van't Hoff equation applied to solid-liquid equilibrium is introduced after which we describe how the temperature and composition dependence of thermodynamic quantities in

the van't Hoff equation can be approximated using a 1st-order Taylor series expansion. Then, we apply our new approach to predict the aqueous solubility of α-glycine using thermodynamic data reported in literature at 298.15 K and assess prediction sensitivity to modelling approach and data quality. Finally, the approach is used to estimate the aqueous solubility of β and γ-glycine using ancillary thermodynamic data reported in literature.

3.2 The van't Hoff Equation

We consider a solute (denoted as 1) which is dissolved in a solvent (denotes as 2). At sufficiently high concentrations, the solute precipitates as a pure solid. The temperature dependence of the mole fraction of the solute $x_{1,\text{sat}}(T)$ at saturation can be expressed by the so-called 'van't Hoff' equation:

$$\frac{d \ln x_{1,\text{sat}}(T)}{dT} = \frac{1}{RT^2}\left[\frac{\bar{h}_1^l(T, x_{1,\text{sat}}(T)) - h_1^{s,\circ}(T)}{1-(1-x_{1,\text{sat}}(T))\partial \ln \gamma_2(T, x_{1,\text{sat}}(T))/\partial x_1}\right] \tag{3.1}$$

$$= \frac{1}{RT^2}\left[\frac{a(T, x_{1,\text{sat}}(T))}{1-b(T, x_{1,\text{sat}}(T))}\right] \tag{3.2}$$

where R is the gas constant, $\bar{h}_1^l(T, x_{1,\text{sat}}(T))$ is the partial molar enthalpy of the solute in a saturated solution at temperature T, and $h_1^{s,\circ}(T)$ is the molar enthalpy of the pure solid at T (which is not a function of solution composition). The solute activity coefficient γ_1 is defined with the chemical potential of the solute written as $\beta\mu_1^l = \beta\mu_1^{\circ,l}(T) + \ln x_1\gamma_1$, where $\mu_1^{\circ,l}(T)$ is the chemical potential of pure liquid solute at temperature T. The quantity $\bar{h}_1^l(T, x_{1,\text{sat}}(T)) - h_1^{s,\circ}(T)$ 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 $\beta\mu_2^l = \beta\mu_2^{\circ,l}(T) + \ln x_2\gamma_2$, where $\mu_2^{\circ,l}(T)$ is the chemical potential of pure solvent at temperature T) with composition, evaluated at saturation temperature and composition. For convenience, we refer to $\bar{h}_1^l(T, x_{1,\text{sat}}(T)) - h_1^{s,\circ}(T)$ 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, it is possible to estimate both a and b from experimentally measurable thermodynamic quantities and, if they were known in addition to a single solubility point, the solubility at any temperature could be predicted through integration of Eq. (3.1).

Expressions equivalent to Eq. (3.1) have been reported by various authors [72, 73, 74, 75]; 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):

$$\beta\mu_1^{\circ,s}(T) = \beta\mu_1^l(T, x_{1,\text{sat}}(T)) \tag{3.3}$$

where $\beta = (RT)^{-1}$, R is the gas constant, $\mu_1^{\circ,s}(T)$ is the chemical potential of pure solid solute and $\mu_1^l(T, x_{1,\text{sat}}(T))$ is the chemical potential of the solute in a saturated solution. Equation (3.3)

is valid at any point along the solubility curve. More generally, small changes in the chemical potential terms can be expressed as:

$$d\beta\mu_1^{\circ,s}(T) = \frac{\partial\beta\mu_1^{\circ,s}(T)}{\partial\beta}d\beta, \tag{3.4}$$

$$d\beta\mu_1^{l}(T,x_1) = \frac{\partial\beta\mu_1^{l}(T,x_1)}{\partial\beta}d\beta + \frac{\partial\beta\mu_1^{l}(T,x_1)}{\partial \ln x_1}d\ln x_1, \tag{3.5}$$

which results from the fact that they are continuous functions of T and x_1. Combining Eq. (3.4) and (3.5), while noting that Eq. (3.3) can be expressed as $d\beta\mu_1^{\circ,s}(T) = d\beta\mu_1^{\circ,l}(T, x_{1,sat}(T))$, gives:

$$\frac{\partial\beta\mu_1^{\circ,s}(T)}{\partial\beta}d\beta = \frac{\partial\beta\mu_1^{l}(T,x_{1,\text{sat}}(T))}{\partial\beta}d\beta + \frac{\partial\beta\mu_1^{l}(T,x_{1,\text{sat}}(T))}{\partial \ln x_1}d\ln x_{1,\text{sat}}(T) \tag{3.6}$$

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

$$\frac{d\ln x_{1,\text{sat}}(T)}{d\beta} = -\left[\frac{\bar{h}_1^{l}(T,x_{1,\text{sat}}(T)) - h_1^{\circ,s}(T)}{\partial\beta\mu_1^{l}(T,x_{1,sat}(T))/\partial\ln x_1}\right], \tag{3.7}$$

where we make use of the Gibbs-Helmholtz relation $\bar{h}_i = \partial\beta\mu_i/\partial\beta$ where $\bar{h}_i$ is the partial molar enthalpy of species i. The denominator in Eq. (3.7) can be expressed as:

$$\frac{\partial\beta\mu_1^{l}(T,x_{1,\text{sat}}(T))}{\partial\ln x_1} = \left[1 + \frac{\partial\ln\gamma_1(T,x_{1,\text{sat}}(T))}{\partial\ln x_1}\right], \tag{3.8}$$

which, coupled with Eq (3.7):

$$\frac{d\ln x_{1,\text{sat}}(T)}{dT} = \frac{1}{RT^2}\left[\frac{\bar{h}_1^{l}(T,x_{1,\text{sat}}(T)) - h_1^{\circ,s}(T)}{1+\partial\ln\gamma_1(T,x_{1,\text{sat}}(T))/\partial\ln x_1}\right] \tag{3.9}$$

which is an expression identical to that reported previously [72, 73, 74, 75]. Since data relating to the activity of solvent in solution is more widely reported, we use the Gibbs-Duhem relationship to replace the solute activity coefficient term:

$$\frac{\partial\ln\gamma_1}{\partial\ln x_1} = -\frac{x_2}{x_1}\frac{\partial\ln\gamma_2}{\partial\ln x_1} = -x_2\frac{\partial\ln\gamma_2}{\partial x_1} \tag{3.10}$$

noting that we make the arbitrary choice to use the partial derivative with respect to x_1, rather than $\ln x_1$ which, when combined with Eq. (3.9) gives Eq. (3.1). For interested readers, we report another derivation in our Appendix A, based entirely on earlier work found in [74]. As an aside, Eq. (3.1) is sometimes reported in the form:

$$\frac{d\ln x_1}{dT} = \frac{\Delta H^{vH}}{RT^2}, \tag{3.11}$$

where ΔH^{vH} is referred to as the "van't Hoff enthalpy of solution" [75, 76].

The utility of Eq. (3.1) is two-fold. First, if the solubility of a compound is known as a function of temperature, the right side can be determined leading to an estimate of the differential heat of solution and related solution properties [74, 77]. 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 [78].

In general, a and b in Eq. (3.2) are functions of both temperature and composition. 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.

3.2.1 An Approximation for Quantities in the van't Hoff Equation

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 mathematical technique used to construct locally accurate but increasingly approximate models of analytic, multi-variate functions in terms of their partial derivatives evaluated at a chosen reference point.

For convenience, we apply the 1$^{\text{st}}$-order approximation of Taylor's theorem which, for a two-variable continuous function, is given as:

$$f(x,y) \approx f(x_0,y_0) + \frac{\partial f(x_0,y_0)}{\partial x}(x-x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y-y_0) \tag{3.12}$$

$$\approx f(x_0,y_0) + x_0\frac{\partial f(x_0,y_0)}{\partial x}\left(\frac{x}{x_0}-1\right) + y_0\frac{\partial f(x_0,y_0)}{\partial y}\left(\frac{y}{y_0}-1\right). \tag{3.13}$$

It should be noted that the 1$^{\text{st}}$-order expansion will limit the accuracy of approximation far from the chosen reference point. In this chapter, we perform the following expansions for a and b, respectively:

$$a(T,x_1) \approx a_0(T_0,x_{1,0}) + a_1\left(\frac{T}{T_0}-1\right) + a_2\left(\frac{x_1}{x_{1,0}}-1\right) \tag{3.14}$$

$$b(\beta,x_1) \approx b_0(\beta_0,x_{1,0}) + b_1\left(\frac{T_0}{T}-1\right) + b_2\left(\frac{x_1}{x_{1,0}}-1\right), \tag{3.15}$$

where definitions for model coefficients and their interpretation are presented in Table 3.1 and nomenclature is defined in Table 3.2.

Approximating $a(T,x_1)$ in terms of T and x_1 which makes it possible to use heat capacity data to evaluate a_1 since $\partial h/\partial T = c_p$. While approximating $b(\beta,x)$ in terms of β and x makes it possible 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. (3.14) and (3.15) into Eq. (3.2), the solubility relation becomes:

$$\frac{d\ln x_{1,\text{sat}}(T)}{dT} \approx \frac{1}{RT^2}\left[\frac{a_0 + a_1(T/T_0-1) + a_2(x_{1,\text{sat}}(T)/x_{1,\text{sat}}(T_0)-1)}{1-(b_0+b_1(T_0/T-1)+b_2(x_{1,\text{sat}}(T)/x_{1,\text{sat}}(T_0)-1))}\right], \tag{3.16}$$

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.

Table 3.1: Taylor series expansion coefficient definitions

Coefficient	Definition
a_0	$(\bar{h_1}^{l}(T_0, x_{1,\text{sat}}(T_0)) - h_1^{s,\circ}(T_0))$
a_1	$T_0(\bar{c}_{p,1}^{l}(T_0, x_{1,\text{sat}}(T_0)) - c_{p,1}^{s,\circ}(T_0))$
a_2	$x_{1,\text{sat}}(T_0)(\partial(\bar{h_1}^{l}(T_0, x_{1,\text{sat}}(T_0)) - h_1^{s,\circ}(T_0))/\partial x_1)$
b_0	$(x_{2,\text{sat}}(T_0)\partial \ln \gamma_2(T_0, x_{1,\text{sat}}(T_0))/\partial x_1)$
b_1	$-\beta_0 \left[x_{1,\text{sat}}(T_0)(\partial(\bar{h_1}^{l}(T_0, x_{1,\text{sat}}(T_0)) - h_1^{s,\circ}(T_0))/\partial x_1)\right]$
b_2	$x_{1,\text{sat}}(T_0)(-\partial \ln \gamma_2(T_0, x_{1,\text{sat}}(T_0))/\partial x_1 + x_{2,\text{sat}}(T_0)\partial^2 \ln \gamma_2(x_{1,\text{sat}}(T_0))/\partial x_1^2)$

Table 3.2: Thermodynamic property definitions

Property	Name	Definition
$\bar{h}_1^l$	Solute partial molar enthalpy (solution)	$\bar{h}_1^l = h_1^{\infty,l} + \frac{\partial \ln \gamma_1^*}{\partial \beta}$
$h_1^{\infty,l}$	Solute partial molar enthalpy in solution at infinite dilution	
$h_1^{s,\circ}$	Pure solute molar enthalpy (solid)	$h_1^{s,\circ} = h_1^{\infty,l} - \Delta h^{\infty,s}$
$\Delta h^{\infty,s}$	Enthalpy of solution (to infinite dilution)	
$\bar{c}_{p,1}^l$	Solute partial molar heat capacity (solution)	$\bar{c}_{p,1}^l = c_{p,1}^{l,\circ} + \frac{\partial^2 \ln \gamma_1}{\partial T \partial \beta}$
$c_{p,1}^{s,\circ}$	Pure solute molar heat capacity (solid)	
$\ln \gamma_i$	Activity coefficient for component i (Raoult's Law)	
$\ln \gamma_i^*$	Activity coefficient for component i (Henry's Law)	$\ln \gamma_i^* = \ln \frac{\gamma_i}{\gamma_i^\infty}$
$\ln \gamma_i^\infty$	Activity coefficient at infinite dilution	$\lim_{x_i \to 0} \ln \gamma_i$

For notational convenience, the composition dependence of coefficients in Eq. (3.16) will be represented by x_1 and $x_{1,0}$ in place of $x_{1,\mathrm{sat}}(T)$ and $x_{1,\mathrm{sat}}(T_0)$ respectively throughout the remainder of the chapter. However, it should be emphasized that these refer to the quantities evaluated at or along the solubility curve.

3.3 Application to the Aqueous Solubility of α, β and γ-Glycine

Glycine is the simplest amino acid [79, 80], known to crystallize in one of three polymorphs identified as γ, α, and β (in order of decreasing thermodynamic stability at ambient conditions [79, 81]). It's chemical structure is illustrated in Fig. 3.1. Different crystallization techniques (i.e. cooling,

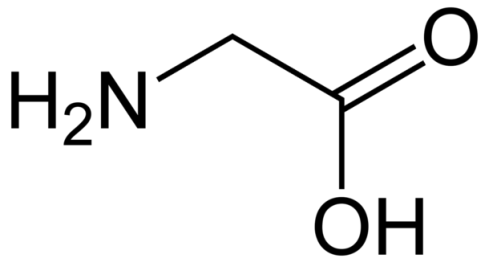


Figure 3.1: Glycine molecular structure.

evaporation, antisolvent) 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 [81]. 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. [80] and [82] for reviews). As pointed out in recent works [81, 83], aqueous solubility data for α and γ-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, γ-glycine, and even for α-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 β-glycine, there is currently limited solubility data available, most likely due to rapid re-crystallisation to α-glycine [84].

Measurements of thermodynamic properties other than solubility (e.g., enthalpy of solution, vapor pressure, etc.) exist over a fairly broad range of temperatures [80] for the glycine-water system; however, the 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. For glycine-water mixtures, the Pitzer [85], mean spherical approximation (MSA) [85], modified Wilson [86], and PC-SAFT [87, 66], as well as many other, models have been used successfully to describe the liquid phase.

Given the reported uncertainty with glycine polymorph temperature dependent solubility and the availability of thermodynamic data other than solubility reported at 298.15 K, it is possible to make predictions for temperature dependent solubility using Eq. (3.1). In this section, the application of Eq. (3.1) to the estimation of glycine polymorph aqueous solubility is presented, providing estimates from their eutectic point to the solid-solution-vapor triple point (i.e. intersection of solubility and boiling curve).

The remainder of this section is structured as follows. First, a review of glycine polymorph aqueous solubility data reported in the literature is presented. Then, the estimation of coefficients in Eq. (3.1) is described and estimates based on available experimental data provided. Following this, solubility predictions based on Eq. (3.1) for α-glycine are shown and 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 is discussed. Finally, predictions for the aqueous solubility of β and γ-glycine are provided.

3.3.1 Solubility Review

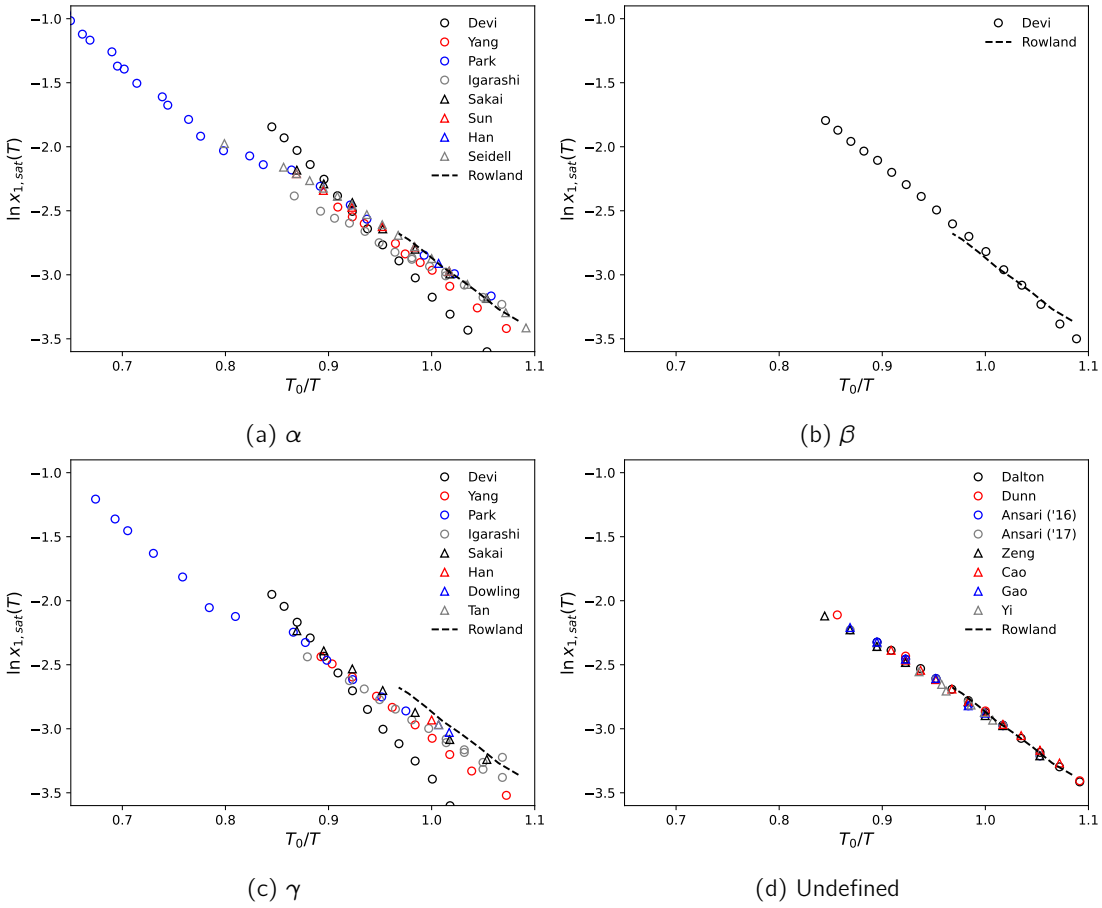
There are numerous reports of glycine-water solubility measurements in the literature. Some articles present data for an identified polymorph (see [88] for α, β and γ, [89, 12, 90, 91, 92] for α and γ, [93, 80, 94] for α and [95] for γ), while others refer only to "glycine" (see [96, 97, 98, 99, 100, 101, 102, 25]) — 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 [80, 82]; however, both reviews have specific areas of focus and limitations. Datta et al. [82], 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 [80] 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 α-glycine, which includes data points reported as α-glycine as well as data where the glycine polymorph is unspecified; however, β and γ-glycine were out of the scope of the work.

The primary aim of this glycine-water solubility review is to consider other glycine polymorphs alongside α-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. α, β or γ) it was included. Then, if the solubility data was for an undefined glycine polymorph, it was included only if the dataset included data above 330 K. Data meeting either of these criteria are presented in Fig. 3.2.

We note that some solubility compilations (e.g., see Table A.5 in Ref. [103]) show glycine solubility up to 373 K, without providing a source reference. However, we believe this is based on an extrapolation of lower temperature data as reported early on [96]; this was also before polymorphism of glycine was established. From Fig. 3.2, each data set shows the solubility of



(a) α (b) β

(c) γ (d) Undefined

Figure 3.2: "Van't Hoff" style plot of literature solubility data [88, 89, 12, 93, 90, 91, 92, 94, 95, 96, 97, 98, 99, 100, 101, 102, 25] for α-, β-, γ-, and undefined glycine between 270–430 K. Note that $T_0 = 298.15$ K. The black dashed line is the α-glycine solubility estimate from Ref. [80] and is shown in all plots as a guide.

glycine in water increasing with temperature as generally expected. For data labeled as either α or γ-glycine, there are a similar number of datasets 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 α and γ-glycine respectively. Interestingly, the labeled polymorph datasets have

greater variability, while the undefined polymorph datasets 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 α-glycine. Based on the slope of their data, we observe that data from Devi and Igarashi for both α and γ-glycine is inconsistent with each other and other data reported in the literature.

Based on our review, there is only one report of β-glycine in the literature; however, it is interesting to note that, as mentioned previously, research suggest direct measurement of β-glycine in pure water are challenging due to kinetic instability in water and resulting rapid transformation to α-glycine [84, 104]. Taking Rowland's estimate for α-glycine as a guide, it appears data reported as β-glycine could be α-glycine.

3.3.2 Methodology

It should be noted that there are many ways in which experimental data could be used to evaluate the coefficients in Eq. (3.16), resulting from the many relationships between thermodynamic quantities. In this chapter, we relate coefficients to experimental measurements performed at a single temperature. This is advantageous in the sense that it limits the amount of thermodynamic data required to make solubility predictions. In the following subsection, an interpretation for each a and b coefficients in the context of available experimental data reported for the glycine-water system is presented. After which, each quantity is evaluated.

Determination of a coefficients

The a coefficient can be expanded as

$$\begin{aligned} a(T, x_{1,\text{sat}}(T)) \approx{}& \Delta\bar{h}_1^{(0)} \\ &+ T_0 \frac{\partial \Delta\bar{h}_1^{(0)}}{\partial T}\left(\frac{T}{T_0} - 1\right) \\ &+ x_{1,\text{sat}}(T_0)\frac{\partial \Delta\bar{h}_1^{(0)}}{\partial x_1}\left(\frac{x_{1,\text{sat}}(T)}{x_{1,\text{sat}}(T_0)} - 1\right) \end{aligned} \tag{3.17}$$

where $\Delta\bar{h}_1^{(0)} = \bar{h}_1^{l}(T_0, x_{1,\text{sat}}(T_0)) - h_1^{s,\circ}(T_0)$ could be evaluated directly if "differential heat of solution" measurements were available; however, in the absence of these measurements, it can be shown that:

$$\Delta\bar{h}_1^{(0)} = \Delta h^{\infty,s}(T_0) + \frac{\partial \ln \gamma_1^{*,(0)}}{\partial \beta} \tag{3.18}$$

where $\Delta h^{\infty,s}$ is the enthalpy of solution when forming an "infinitely dilute" solution [105]. It should be noted that we introduce $\gamma_1^* = \gamma_1/\gamma_1^\infty$ as the solute activity coefficient defined relative to an infinitely dilute solution, where $\gamma_1^\infty = \lim_{x_1 \to 0} \gamma_1$. The quantity $\partial \ln \gamma_1^{*,(0)}/\partial\beta = \partial \ln \gamma_1^*(T_0, x_{1,\text{sat}}(T_0))/\partial\beta$ 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:

$$\frac{\Delta H_{dil}(T, x^f, x^i)}{n_1} = \left(\sum_i^N \frac{n_i}{n_1}\frac{\partial \ln \gamma_i(T,x)}{\partial \beta}\right)^f - \left(\sum_i^N \frac{n_i}{n_1}\frac{\partial \ln \gamma_i(T,x)}{\partial \beta}\right)^i, \tag{3.19}$$

where superscript i and f denote the solution at initial and final concentration (i.e., before and after dilution), respectively. In addition, since $\Delta h^{\infty,s}(T_0)$ is independent of composition, it can be shown that:

$$\frac{\partial \Delta \bar{h}_1^{(0)}}{\partial x_1} = \frac{\partial^2 \ln \gamma_1^{(0)}}{\partial x_1 \partial \beta}, \tag{3.20}$$

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:

$$\frac{\partial \Delta \bar{h}_1^{(0)}}{\partial T} = \Delta \bar{c}_{p,1}^{(0)} \tag{3.21}$$

where $\Delta \bar{c}_{p,1}^{(0)} = \bar{c}_{p,1}^{l}(T_0, x_{1,\text{sat}}(T_0)) - c_{p,1}^{s,\circ}(T_0)$, $c_{p,1}^{\circ,s}$ is the molar heat capacity of crystalline glycine, which can be measured directly, and $\bar{c}_{p,1}^{l}$ is the partial molar heat capacity of glycine in solution which can be interpreted as:

$$\bar{c}_{p,1}^{l}(T_0, x_{1,\text{sat}}(T_0)) = c_{p,1}^{\circ,l}(T_0) + \frac{\partial^2 \ln \gamma_1(T_0, x_{1,\text{sat}}(T_0))}{\partial T \partial \beta}, \tag{3.22}$$

which, given an excess Gibbs free energy model, can be estimated by regressing solution heat capacity data, since:

$$c_p^{soln}(T,x) = \sum_i^N x_i c_{p,i}^{l,\circ}(T) + \sum_i^N x_i \frac{\partial^2 \ln \gamma_i(T,x)}{\partial T \partial \beta}. \tag{3.23}$$

Determination of b coefficients

In the context of this work:

$$\begin{aligned} b(\beta, x_{1,\text{sat}}(T)) \approx\ & x_{2,\text{sat}}(T_0)\frac{\partial \ln \gamma_2^{(0)}}{\partial x_1} \\ & + \beta_0 \frac{\partial (x_{2,\text{sat}}(T_0)(\partial \ln \gamma_2^{(0)}/\partial x_1))}{\partial \beta}\left(\frac{T_0}{T} - 1\right) \\ & + x_{1,\text{sat}}(T_0)\frac{\partial (x_{2,\text{sat}}(T_0)(\partial \ln \gamma_2^{(0)}/\partial x_1))}{\partial x_1}\left(\frac{x_{1,\text{sat}}(T)}{x_{1,\text{sat}}(T_0)} - 1\right). \end{aligned} \tag{3.24}$$

where again $\ln \gamma_2^{(0)} = \ln \gamma_2(T_0, x_{1,\text{sat}}(T_0))$ is introduced for notational convenience. It can be shown that:

$$\frac{\partial}{\partial x_1}\left(x_{2,\text{sat}}(T_0)\frac{\partial \ln \gamma_2^{(0)}}{\partial x_1}\right) = -\frac{\partial \ln \gamma_2^{(0)}}{\partial x_1} + x_{2,\text{sat}}(T_0)\frac{\partial^2 \ln \gamma_2^{(0)}}{\partial x_1^2}, \tag{3.25}$$

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:

$$\frac{\partial}{\partial\beta}\left(x_{2,\text{sat}}(T_0)\frac{\partial \ln\gamma_2^{(0)}}{\partial x_1}\right) = -x_{1,\text{sat}}(T_0)\frac{\partial^2 \ln\gamma_1^{*,(0)}}{\partial x_1 \partial\beta}, \tag{3.26}$$

which can be evaluated in the same way as Eq. (3.20).

Excess Gibbs Free Energy Models

Several parameters required to estimate the coefficients in Eq. (3.16) are derivatives of continuous thermodynamic quantities. Given the discrete nature of experiments, coupled with the absence of approaches to measure the thermodynamic quantities in Eq (3.1) directly, estimation of required parameters requires correlation of data using models. In the context of thermodynamics, the excess Gibbs free energy relates most of the quantities in Eq. (3.16) to experimental data and as such, a mathematical model describing the excess Gibbs free energy is required. 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 [106].

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

$$\beta G^E = (n_1 v_1^L + n_2 v_2^L)\phi_1\phi_2\chi' \tag{3.27}$$

where v_i^L is the molar volume of component i, ϕ_i is the volume fraction of component i, and χ' is a binary interaction parameter. This gives:

$$\ln\gamma_2 = \beta v_2^L \phi_1^2 \chi' \tag{3.28}$$

$$\ln\gamma_1 = \beta v_1^L \phi_2^2 \chi' \tag{3.29}$$

For convenience, we define $\chi(\beta) = \chi'\beta v_2^L$ and $v^* = v_1^L/v_2^L$ which gives:

$$\ln\gamma_2 = \phi_1^2\chi \tag{3.30}$$

$$\ln\gamma_1 = v^*\phi_2^2\chi, \tag{3.31}$$

where we have assumed χ 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:

$$\phi_1 = \frac{x_1 v^*}{x_1 v^* + x_2} \tag{3.32}$$

$$\phi_2 = \frac{x_2}{x_1 v^* + x_2} = 1 - \phi_1 \tag{3.33}$$

The excess enthalpy of this model, assuming χ' is a function of temperature only and v^* is independent of temperature and composition, is given by:

$$H^E = n_1 v_1^L \phi_1 \phi_2 \chi_1' + n_2 v_2^L \phi_1 \phi_2 \chi_1', \tag{3.34}$$

and the excess heat capacity is given by:

$$C_p^E = n_1 v_1^L \phi_1 \phi_2 \chi_2' + n_2 v_2^L \phi_1 \phi_2 \chi_2', \tag{3.35}$$

where $\chi_1' = \partial\chi'/\partial\beta$, and $\chi_2' = \partial^2\chi/\partial T\partial\beta$.

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:

$$\beta G^E = (n_1 v_1^L + n_2 v_2^L)\phi_1\phi_2\chi' + n_1 \ln\frac{\phi_1}{x_1} + n_2 \ln\frac{\phi_2}{x_2}, \tag{3.36}$$

which gives:

$$\ln\gamma_2 = \phi_1^2\chi + \left(\ln\frac{\phi_2}{x_2} + 1 - \frac{\phi_2}{x_2}\right) \tag{3.37}$$

$$\ln\gamma_1 = v^*\phi_2^2\chi + \left(\ln\frac{\phi_1}{x_1} + 1 - \frac{\phi_1}{x_1}\right) \tag{3.38}$$

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 χ' is a function of temperature only and v^* is a constant independent of temperature and composition, which is an assumption used in this work.

3.3.3 Thermodynamic Properties

As discussed above, Rowland [80] 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. (3.16) 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 298.5 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. (3.16) was found.

Solubility at the chosen reference temperature is a crucial thermodynamic property required to make predictions with the approach presented in the theory section. For α-glycine, we have chosen to take the value reported by Rowland [80] (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.

In this work, we model solution properties separately (i.e. $\ln\gamma_2$ is modeled independently of $\Delta H^{dil}/\mathrm{n}$ and c_p^{soln}), 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 evaluate the coefficients in Eq. (3.16) 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, twelve data sets [107, 108, 109, 87, 110, 111, 85, 112, 113, 114, 115, 116] were found from which the water activity coefficient as a function of solution composition can be derived. For seven of the datasets, the water activity coefficient γ_2 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:

$$\Phi = -\frac{1}{M_2 m_1} \ln(\gamma_2 x_2), \tag{3.39}$$

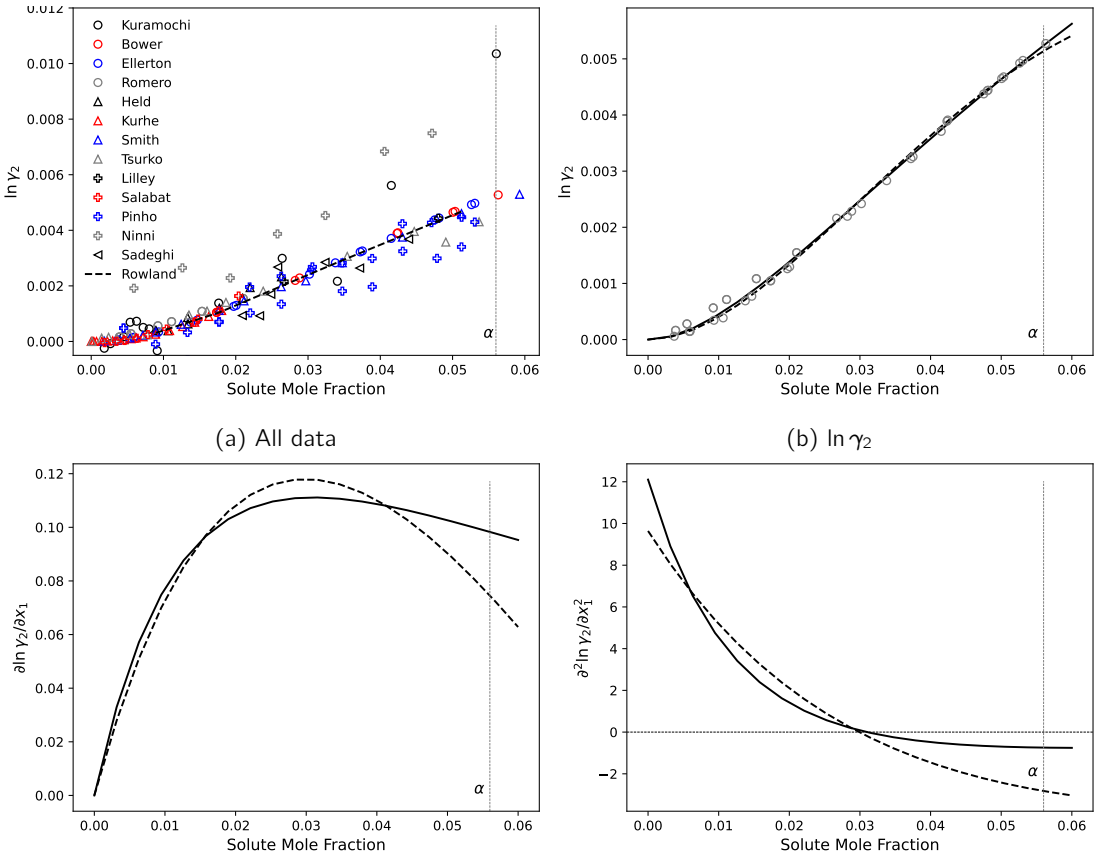
where x_2 is the mole fraction of water, $M_2 = 0.018\,\mathrm{kg\,mol^{-1}}$ is the molecular weight of water, and m_1 is glycine molality. Four data sets reported values for water activity a_2 at various glycine-water solution concentrations (at 298.15 K) which were converted to activity coefficients via:

$$a_2 = \gamma_2 x_2. \tag{3.40}$$

The final dataset 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):

$$P = x_2 \gamma_2 P_2^{sat}(T), \tag{3.41}$$

where P is the pressure, and $P_2^{sat}(T)$ is the saturated vapor pressure of pure water at 298.15 K, taken as the value provided in the dataset. All data, converted to $\ln \gamma_2$ and plotted as a function of solution composition, are given in Fig. 3.3a. From Fig. 3.3a, the general trend indicates that $\ln \gamma_2$ increases with glycine mole fraction, meaning that $\gamma_2 > 1$. Excluding the data from Ninni, the data appears to follow the same trend; however, there is significant variability between datasets, particularly above $x_1 = 0.03$. 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.



(a) All data

(b) $\ln \gamma_2$

(c) $\partial \ln \gamma_2 / \partial x_1$

(d) $\partial^2 \ln \gamma_2 / \partial x_1^2$

Figure 3.3: (a) Literature glycine-water solution activity data at 298.15 K [107, 108, 109, 87, 110, 111, 85, 112, 113, 114, 115, 116], (b) model fits to selected data activity data, (c) 1st-order partial derivative for fitted models with respect to concentration, and (d) 2nd-order partial derivative for fitted models with respect to concentration. Solid lines are the predictions of the Scatchard-Hildebrand model, dashed lines are the predictions of the Scatchard-Hildebrand-Flory-Huggins model, and symbols are converted experimental data from the literature.

From Fig. 3.3b, 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 1st and 2nd partial derivatives and their values at glycine polymorph solubility. For example, the 1st partial derivative evaluated at the estimated α-glycine solubility is 0.098 and 0.075 for the Scatchard-Hildebrand and Scatchard-Hildebrand-Flory-Huggins model, respectively, while the 2nd 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.

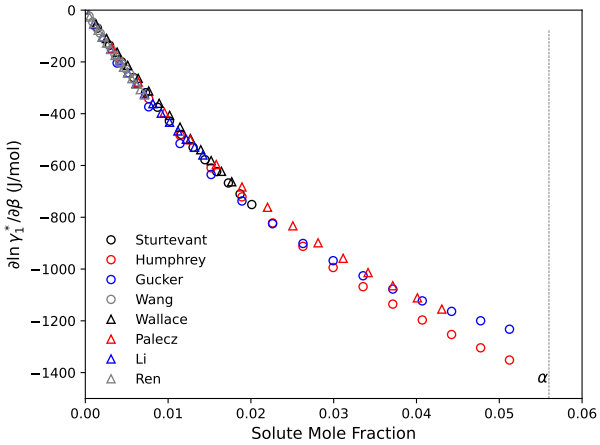
Enthalpy of Dilution Data, Fits and Derived Properties

Again, following Rowland there are at least 8 datasets [117, 118, 119, 120, 121, 122, 123, 124] available in the literature reporting enthalpy of dilution measurements. For the purpose of review, each dataset was correlated by Eq. (3.19) to allow direct comparison, noting that raw measurements cannot be compared graphically. From the best fit to each data set, points corresponding to $\partial \ln \gamma_1^* / \partial \beta$ were generated corresponding to the range over which measurements were reported. The resulting processed data is presented in Fig. 3.4a. From Fig. 3.4a, data generated from best fits to each data set appear consistent from 0–0.02 mole fraction. However, beyond 0.02, there is uncertainty resulting from a lack of data and inconsistency between the data that is available. This is problematic because an estimate of the partial derivative of this curve at the estimated solubility is required to evaluate model parameters and thus extrapolation is necessary—which may introduce uncertainty.

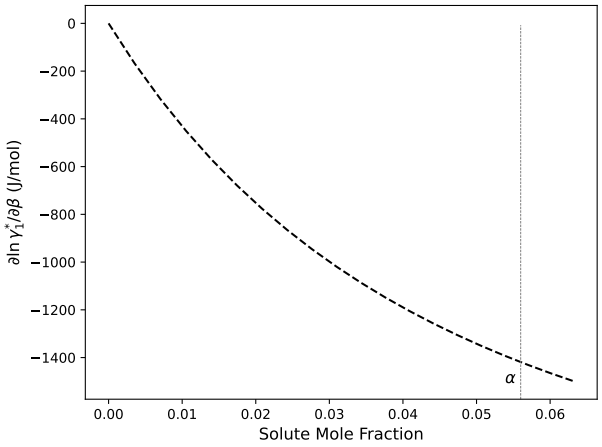
Again, Eq. (3.19) was regressed against the aggregated data set (i.e. the data sets used by Rowland) and the results are presented in Fig. 3.4b, with $\partial^2 \ln \gamma_1^* / \partial x_1 \partial \beta$ in Fig. 3.4c. It should be noted that our best fit estimate for $\partial \ln \gamma_1^* / \partial \beta$ is in excellent agreement with that presented in Ref. [118].

Solution Heat Capacity Data, Fits and Derived Properties

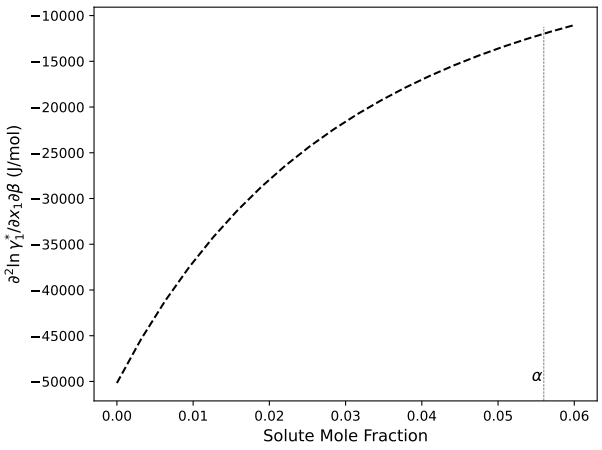
Rowland references 9 data sets reporting measurements relating to the heat capacity of glycine-water solutions. However, only 5 [77, 125, 126, 127, 128] have data for solutions at 298.15 K. These data are presented in Fig. 3.5a. It should be noted that, where data was reported in terms of the "apparent molar heat capacity", data has been converted to solution heat capacity using pure water molar heat capacity of $75.2\,\mathrm{J\,mol^{-1}}$. From Fig. 3.5a, 4 data sets appear consistent between 0 and 0.01, corresponding to the data used by Rowland in their global fit. For the purpose of this work, we use these data sets. It should be highlighted again that data doesn't extend to solution compositions required to accurately estimate properties at estimated solubility. In addition, there is only one dataset that extends beyond a solution composition of 0.01. Thus, we



(a) $\partial \ln \gamma_1^* / \partial \beta$ (Converted Data)

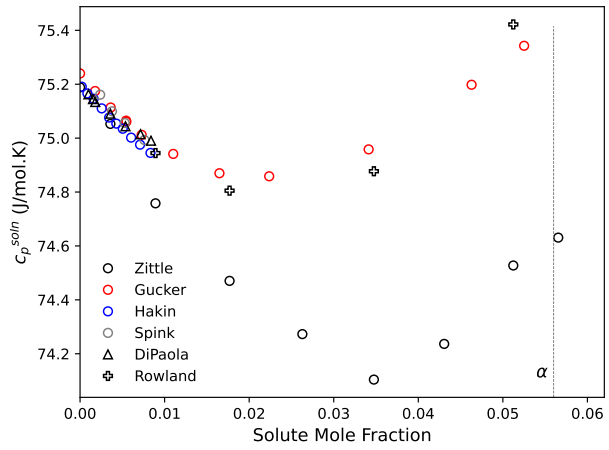


(b) $\partial \ln \gamma_1^* / \partial \beta$ (Best Fit)

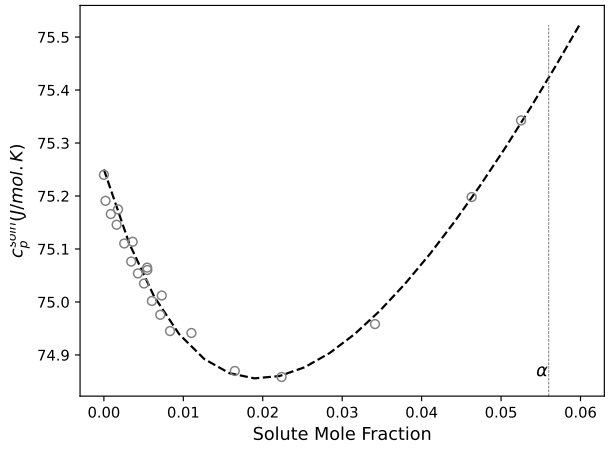


(c) $\partial^2 \ln \gamma_1^* / \partial x_1 \partial \beta$ (Best Fit)

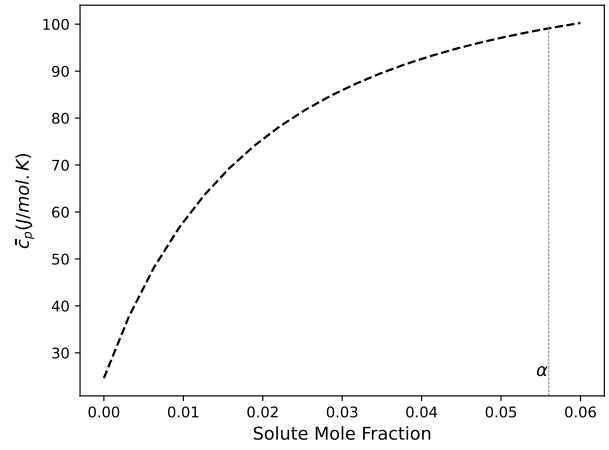
Figure 3.4: (a) Literature glycine-water solution enthalpy of dilution processed data at 298.15 K, (b) and (c) $\partial \ln \gamma_1^*$ and $\partial^2 \ln \gamma_1^* / \partial x_1 \partial \beta$ derived from Scatchard-Hildebrand best fit to data in (a). Open symbols are values derived from fits to literature data [117, 118, 119, 120, 121, 122, 123, 124], grey dashed vertical lines indicate α-glycine solubility at 298.15 K



(a) All data



(b) c_p^{soln}



(c) $\bar{c}_{p,1}$

Figure 3.5: (a) Glycine-water solution heat capacity data at 298.15 K [77, 125, 126, 127, 128], (b) and (c) c_1^{soln} and $\bar{c}_{p,1}$ derived form Scatchard-Hildebrand best fit to selected data in (a). Grey dashed vertical lines indicate α-glycine aqueous solubility at 298.15 K

have to extrapolate regressed models to evaluate required parameters. The fit and corresponding partial molar heat capacity are given in Figs. 3.5b and 3.5c.

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. (3.16). Both have been reported in the literature and are presented in Table 3.3. We note that both $\Delta h^{\infty,s}$ and $c_{p,298.15}$ has been measured for α, β and γ-glycine respectively. However, we also note that each property has been measured only once and not validated by additional measurements.

Table 3.3: Enthalpy of solution (to infinite dilution) and crystal molar heat capacity data for α, β and γ-glycine at 298.15 K

Polymorph	$\Delta h^{\infty,s}$ (J mol^{-1}) [105]	$c_{p,298.15}$ (J/mol.K) [129]
α	14523 $\pm$ 76	99.23
β	14198 $\pm$ 73	98.69
γ	14791 $\pm$ 84	96.00

3.3.4 α-Glycine Solubility Predictions

The solubility of α-glycine was predicted as a function of temperature by numerically integrating Eq. (3.16) with the coefficient values presented in Table 3.4, and predictions are shown in Fig. 3.6. It should be noted that the solubility was also estimated with a_0 and b_0 only, which is referred to as the "0$^{\text{th}}$-order" prediction, while predictions based on Eq. (3.16) are referred to as "1$^{\text{st}}$-order" predictions.

From Fig. 3.6, the 0$^{\text{th}}$ and 1$^{\text{st}}$-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 0$^{\text{th}}$ and 1$^{\text{st}}$-order approaches show good agreement in terms of solubility; however, beyond this, predictions diverge. Although solubility data is limited above 340 K, the 1$^{\text{st}}$-order corrections shift solubility predictions toward available measurements, potentially indicating that the 1$^{\text{st}}$-order corrections improve those made by the 0$^{\text{th}}$-order model.

We note that our solubility predictions are consistent with the majority of available direct measurements, as shown in Fig. 3.6. However, our predictions vary considerably compared to Refs. [88] and [90], 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

Table 3.4: Taylor series expansion coefficient values (α, β and γ-glycine)

Coefficient	Scatchard-Hildebrand	Exp. Scatchard-Hildebrand
	α	
a_0 (J/mol)	13103.5	13103.5
a_1 (J/mol)	-96.8	-96.8
a_2 (J/mol)	-671.1	-671.1
b_0 (-)	0.093	0.07
b_1 (-)	0.27	0.27
b_2 (-)	-0.045	-0.15
	β	
a_0 (J/mol)	12597.	12605.
a_1 (J/mol)	1314.5	1261.0
a_2 (J/mol)	-622.2	-625.4
b_0 (-)	0.079	0.019
b_1 (-)	0.251	0.252
b_2 (-)	-0.056	-0.240
	γ	
a_0 (J/mol)	13421.5	13421.5
a_1 (J/mol)	478.8	478.8
a_2 (J/mol)	-677.8	-677.8
b_0 (-)	0.096	0.081
b_1 (-)	0.273	0.273
b_2 (-)	-0.041	-0.131

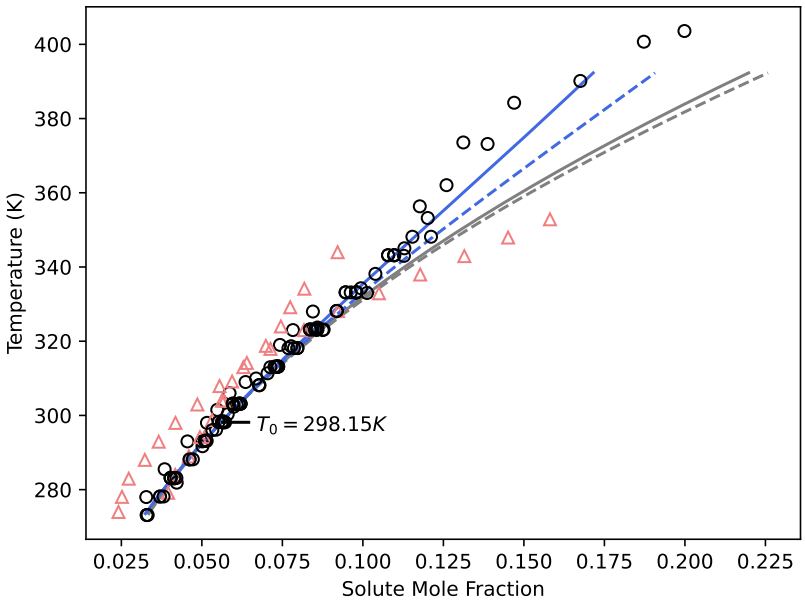


Figure 3.6: Temperature dependent α-glycine solubility predictions for various data modelling approaches (grey lines; 0^{th}-order predictions, blue lines; 1^{st}-order predictions, dashed lines; Scatchard-Hildebrand predictions, solid lines; Scatchard-Hildebrand-Flory-Huggins predictions, black open symbols; direct measurements labeled as α-glycine [89, 12, 93, 91, 92] and undefined glycine [96, 97, 98, 99, 100, 101, 102, 25] Refs. [88] and [90] highlighted with red triangles due to being inconsistent with rest of available 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. (3.16) is shown in Fig. 3.6. For the 0^{th}-order predictions, model impact is limited—becoming significant around 380 K. However, the impact of model selection is greater for the 1^{st}-order predictions. The corrected models begin to diverge at approximately 340 K. At 400 K, the solubility is predicted as 0.17 and 0.19 (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 is 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.

Based on 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—all solubility data above 340 K comes from the same source in which solubility was determined using a novel application of "Differential Scanning Calorimetry (DSC)" [12], 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 and Eq. (3.16), respectively.

Sensitivity Analysis

As shown above, the choice of model used to estimate coefficients in Eq. (3.16) 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. $\ln \gamma_2$, $\Delta H^{dil}/n$ and c_p^{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 $y_{blur} = y_{original} + y_{original}\epsilon$ where y denotes a thermodynamic quantity and ϵ is a

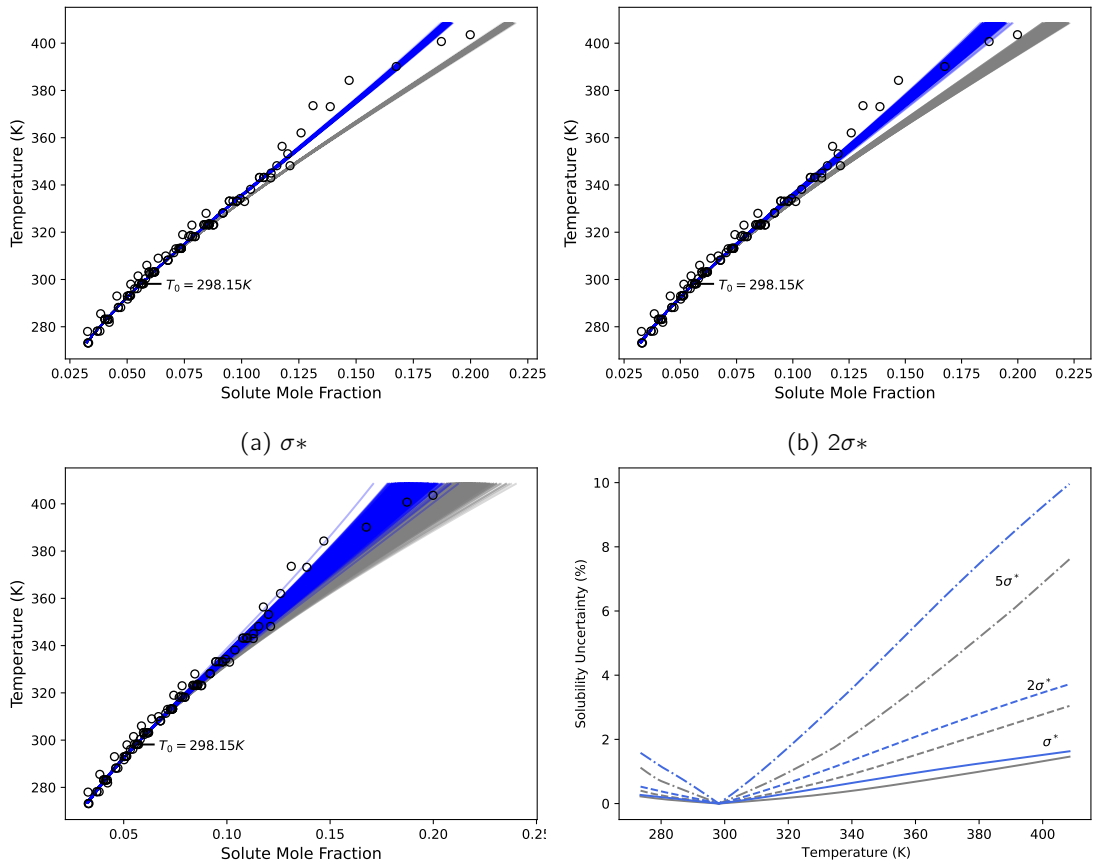
random value drawn from a Gaussian distribution with mean $\mu^* = 0$, and standard deviation, σ^*.

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 Fig. 3.7a. To investigate the effect of uncertainty magnitude, we perform sensitivity analysis at three levels for each thermodynamic quantity; σ^*, $2\sigma^*$ and $5\sigma^*$, where σ^* is defined as 0.01, 0.01 and 0.0001 for $\ln \gamma_w$, ΔH^{dil} and c_p^{soln}, respectively. It should be noted that the chosen σ^* values correspond to approximate uncertainty observed in aggregated data sets shown in Figs. 3.3b, 3.4a and 3.5b. From the available data, measurements of c_p^{soln} are more precise, in terms of % error, when compared to other thermodynamic quantities. Uncertainty levels translated to percentage errors are provided in Table 3.5. Resulting fits to experimental data are presented in Appendix A (see Sensitivity Analysis). From Fig. 3.7, solubility

Table 3.5: Levels used in α-glycine sensitivity analysis

	Quantity	Error (%)
σ^*	$\ln \gamma_2$	1
	ΔH^{dil}	1
	c_p^{soln}	0.01
$2\ \sigma^*$	$\ln \gamma_2$	2
	ΔH^{dil}	2
	c_p^{soln}	0.02
$5\ \sigma^*$	$\ln \gamma_2$	5
	ΔH^{dil}	5
	c_p^{soln}	0.05

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 grey 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 levels. As the level of uncertainty used to "blur" data increases, the uncertainty of predictions far from the reference temperature increases, which is generally 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 σ^* and $2\sigma^*$ uncertainty levels (Figs. 3.7a and 3.7b). This suggests that model selection has an impact, independent of measurement uncertainty, on the validity of the method and motivates



(a) $\sigma*$

(b) $2\sigma*$

(c) $5\sigma*$

(d) % error in x vs. T

Figure 3.7: (a)–(c) Solubility predictions based on excess Gibbs free energy model parameters fit to "simulated" data sets. The grey 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. Grey 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 α-glycine [89, 12, 93, 91, 92] and undefined glycine [96, 97, 98, 99, 100, 101, 102, 25]

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 analysed. From Fig. 3.7d, 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 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 σ^*, $2\sigma^*$, and $5\sigma^*$, respectively.

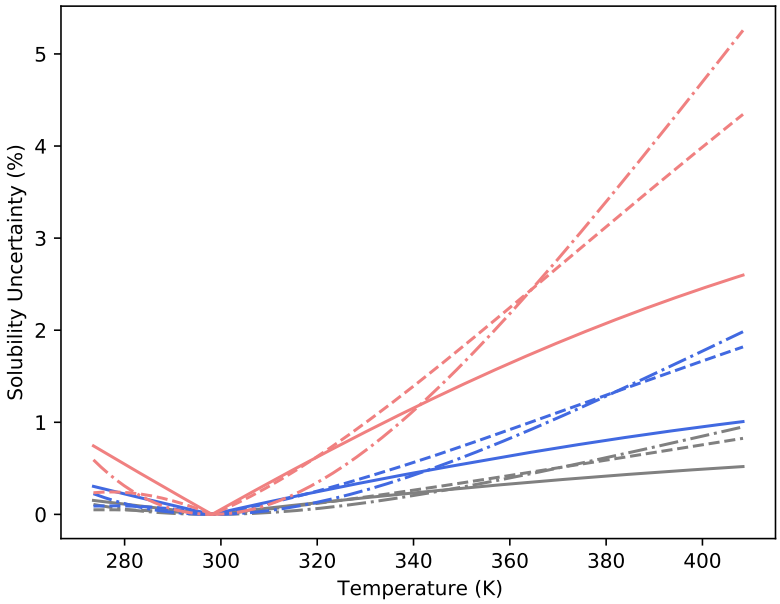


Figure 3.8: Uncertainty of the predicted α-glycine solubility based on the Scatchard-Hildebrand model with fluctuations of σ^* (gray), $2\sigma^*$ (blue), and $5\sigma^*$ (red) in the value of $\ln\gamma_2$ (solid lines), ΔH^{dil} (dashed lines), and c_p^{soln} (dashed-dotted lines).

The sensitivity of solubility predictions (based on the Scatchard-Hildebrand excess Gibbs free energy model) to independent thermodynamic data is shown in Fig. 3.8. We note that, for the case where uncertainty in data is estimated by "blurring" data points and generating new model fits, solubility predictions far from the reference temperature are impacted by: c_p^{soln}, $\Delta H^{dil}/n_1$ and $\ln\gamma_2$ in order of most to least impact, where a greater impact corresponds to larger percentage error. If we consider that the level of uncertainty in c_p^{soln} is 100 times lower than ΔH^{dil} and $\ln\gamma_2$, it suggests that, for accurate solubility predictions, precise measurements of c_p^{soln} are required (i.e. $\pm 0.1\,\mathrm{J\,mol^{-1}\,K^{-1}}$). However, it should be noted that, for the same level of uncertainty, both c_p^{soln} and ΔH^{dil} have a similar impact, while for all levels, uncertainty in $\ln\gamma_2$ has comparatively lower impact. For example, at the $5\sigma^*$ level, the percentage uncertainty in predicted solubility at

400 K is ≈ 2.5, 4 and 5% for $\ln\gamma_2$, ΔH^{dil}, and c_p^{soln}, respectively.

3.3.5 β and γ-Glycine Solubility Predictions

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. 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 $\Delta h^{\infty,\circ}$ and c_p° are available (Table 3.3), alongside regressed excess Gibbs free energy models describing the thermodynamic properties of glycine-water solutions. However, for β-glycine, there is only one report of solubility data available, while for γ-glycine the solubility data is scattered (Fig. 3.2). 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 3.6.

In Table 3.6, $\Delta\mu^{x-y}$ is an estimate of the polymorph free energy difference at 298.15 K, $(c_x/c_y)_{(310\,K)}$ is polymorph solubility ratios measured in water-antisolvent systems with various solvent compositions at 310 K, T_E and x_E are eutectic temperature and composition measurements, x_γ corresponds to direct solubility measurements of γ-glycine, x_α/x_γ corresponds to α/γ solubility ratios derived from direct solubility measurements reported together and $T^{\alpha-\gamma}$ corresponds to estimates for the temperature at which the relative stability of α and γ-glycine changes.

Table 3.6: Summary of data used to estimate β and γ-glycine reference solubility

β	γ
$\Delta\mu^{\gamma-\beta}_{(298.15\,K)}$	$\Delta\mu^{\alpha-\gamma}_{(298.15\,K)}$
$(c_\beta/c_\alpha)_{(310\,K)}$	$(c_\gamma/c_\alpha)_{(310\,K)}$
T_E and x_E	T_E and x_E
—	x_γ
—	x_γ/x_α
—	$T^{\alpha-\gamma}$

Eutectic Temperature and Composition

The eutectic temperatures for α-, β-, and γ-glycine in water have been reported by various authors and recently reviewed [130]. From the available data (see Appendix A (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 $\pm 0.1°C$. The resulting eutectic temperatures are presented alongside selected freezing point measurements in Fig. 3.9. By evaluating the solution composition at

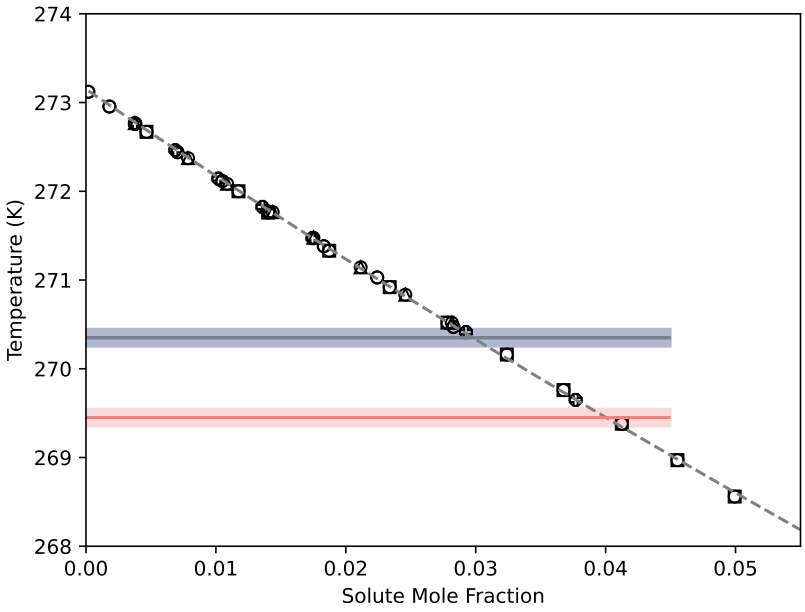


Figure 3.9: Expected eutectic temperatures and freezing curve The open black symbols are freezing curve measurements [131, 132, 133, 134, 135], and the grey dashed line is the best fit line. The solid lines are the estimated eutectic temperature, and the opaque bands indicate the uncertainty. Blue refers α-, grey refers to γ-, and red refers to β-glycine.

which the eutectic temperature range for β- and γ-glycine intersects the freezing point line, we estimate their eutectic composition; giving $x_{E,\beta} = 0.040 \pm 0.001$ and $x_{E,\gamma} = 0.03 \pm 0.001$ 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 asses the resulting performance in relation 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 $x_\gamma/x_\alpha = 0.974 \pm 0.012$. Measurements of β-glycine in pure water are not reported due to rapid re-crystallization 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 Appendix A (Bouchard Ratios)). We find that this approach gives a reasonable estimate for γ/α, thus apply it to β/γ. The estimated solubility ratio was found to be $x_\beta/x_\alpha = 1.13 \pm 0.08$.

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 [136, 129, 105]. The free energy differences are reported as $\Delta\mu^{\alpha,\gamma} = 157 \pm 145\ \mathrm{J\,mol^{-1}\,K^{-1}}$ and $\Delta\mu^{\beta,\gamma} = 277\ \mathrm{J\,mol^{-1}\,K^{-1}}$, giving $\Delta\mu^{\beta,\alpha} = 120\ \mathrm{J\,mol^{-1}\,K^{-1}}$. It should be noted that the reported uncertainty for $\Delta\mu^{\gamma,\alpha}$ 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:

$$\beta(\mu_i^{s,\alpha} - \mu_i^{s,\gamma}) = \ln(x_i^{sat,\alpha}\gamma_i(x_i^{sat,\alpha})) - \ln(x_i^{sat,\gamma}\gamma_i(x_i^{sat,\gamma})) \tag{3.42}$$

$$\approx \ln(x_i^{sat,\alpha}) - \ln(x_i^{sat,\gamma}), \tag{3.43}$$

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. (3.43) we estimate the solubility ratios to be $x_\gamma/x_\alpha = 0.94 \pm 0.05$ and $x_\beta/x_\alpha = 1.05$. It should be noted that the mole fraction solubility ratio estimated for γ/α is in good agreement with those derived from direct measurements (see Appendix A).

Solubility Predictions

Based on 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

difference was used, giving $x_{\gamma,298.15} = 0.052$ (based on $x_{\alpha,298.15} = 0.056$). The resulting solubility, solubility ratio and eutectic point estimates are presented in Figs. 3.10, 3.11 and 3.12, respectively; all of which are based on model paramaters presented in Table 3.4. As expected, the

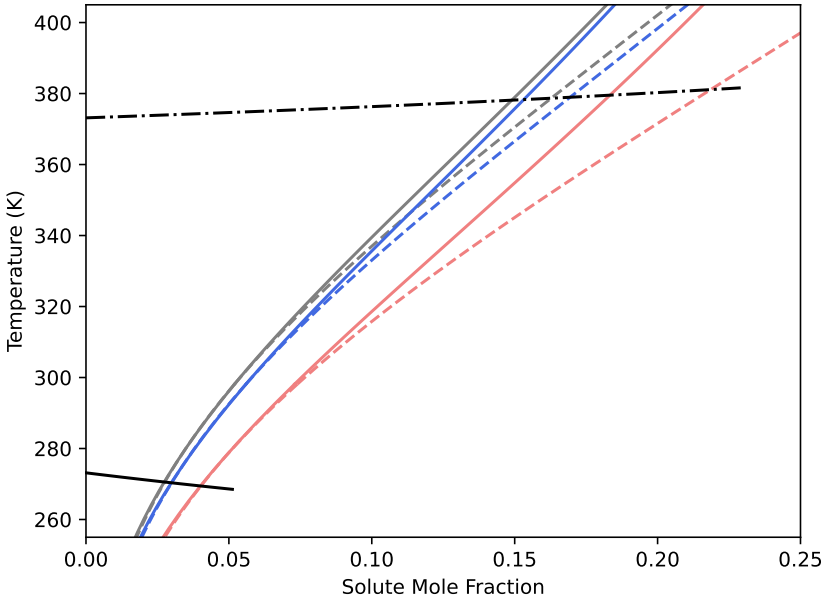


Figure 3.10: Solubility predictions for α, β and γ-glycine based on Eq. (3.16), with coefficients derived from different glycine-water solution models. The dashed lines are the prediction of the Scatchard-Hildebrand model, and the solid lines are the predictions of the Scatchard-Hildebrand-Flory-Huggins model. Blue refers to α-, red refers to β-, and grey refers to γ-glycine solubility in water. The black solid line is the best fit to the freezing point depression curve, and the black dash-dotted line is the ideal boiling curve.

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 Fig. 3.10, 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 Fig. 3.3, which corresponds to diverging estimates for $\partial \ln \gamma_w / \partial x_g$ and $\partial^2 \ln \gamma_w / \partial x_g^2$, 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

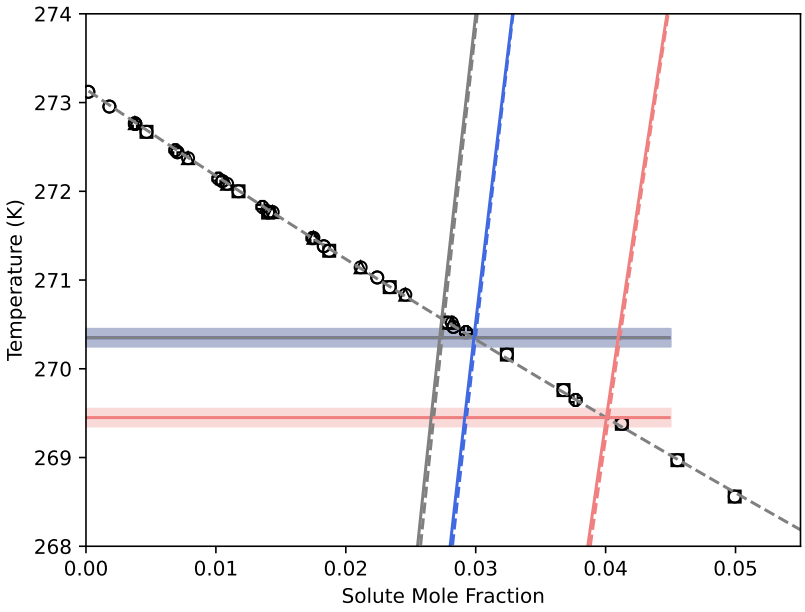


Figure 3.11: Eutectic temperature and composition predictions α-, β-, and γ-glycine based on Eq. (3.16), with coefficients derived from different glycine-water solution models. Open circles are freezing point measurements for water [131, 132, 133, 134, 135], and the grey dashed line is the best-fit line, Blue refers to α-, red refers to β-, and grey refers to γ-glycine. The dashed lines are the predictions of the Scatchard-Hildebrand, and the solid lines are the predictions of the Scatchard-Hildebrand-Flory-Huggins model.

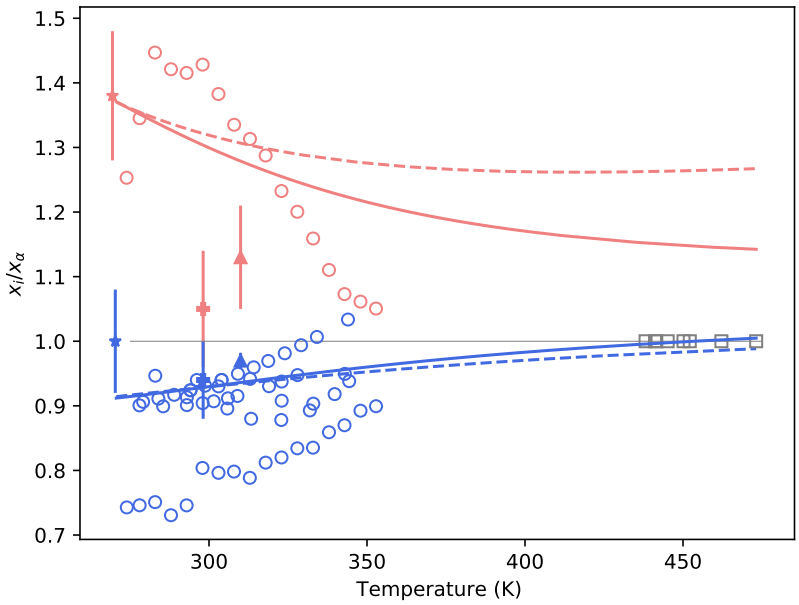


Figure 3.12: Glycine polymorph mole fraction solubility ratio predictions based on Eq. (3.16), with coefficients derived from different glycine-water solution models. Blue refers to the γ-α ratio, and red refers to β-α ratio. The dashed lines is the prediction of the Scatchard-Hildebrand model, and the solid lines Scatchard-Hildebrand-Flory-Huggins model. The open symbols are ratios derived from direct measurements [131, 132, 133, 134, 135] and the pluses are estimates from free energy ratios [129], stars are estimates from eutectic temperatures, triangles are estimates from Ref. [84] and freezing point data, and grey squares are α-γ crossover estimates [137, 105, 138, 12, 139, 140, 141, 142]. Bars indicate uncertainty estimates.

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 to note the Scatchard-Hildebrand-Flory-Huggins model correctly predicts the γ/α 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 β-glycine to be least stable (i.e. highest solubility) in the temperature range explored, agreeing with qualitative observations in the literature. However, predictions for β/α based on the expected eutectic temperature appear to be inconsistent with independent ratio estimates derived from anti-solvent mixture solubility data and free energy difference estimates, though this could be attributed to significant uncertainty associated to both estimates.

Finally, we note that solubility predictions based on different solution models are approximately identical in the eutectic region (as shown in Fig. 3.11), where the predicted eutectic temperature and composition for α-glycine agrees extremely well with data available in literature, as does γ (noting that models for α and γ were not based on the eutectic temperature at all). Our predictions suggests that $-2.7°C$ is an appropriate eutectic temperature for γ-glycine.

3.4 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 α-glycine.

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 β and γ-glycine where previous solubility data is inconsistent or limited, providing estimates for their aqueous solubility between 273–310 K 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. In addition, for compounds that undergo thermal decomposition or polymorph

transition prior to melting.

Chapter 4

A Local Expansion Approximation of Solubility

The results presented in this chapter are in-preparation for submission as a manuscript authored by Andrew Manson, Jan Sefcik and Leo Lue. It should be noted that large sections of this chapter reflect those in the manuscript; however, adaptations have been made to improve readability in the context of this thesis.

4.1 Introduction

In Chapter 3, a predictive approach allowing extrapolation of temperature dependent solubility based on thermodynamic quantities evaluated at a specific point on the solubility curve was proposed. As shown, this approach gave solubility estimates in good agreement with direct measurements for α-glycine in water, around the chosen reference temperature (i.e., 298.15 K). However, estimates were significantly impacted by the approach taken to derive required thermodynamic quantities required to make predictions. Furthermore, the impact of the modeling approach was challenging to quantify given that certainty around α-glycine aqueous solubility provided by direct measurements decreased significantly at temperatures above 340 K.

In this chapter, we propose another approximate model describing solubility as a function of temperature but instead of approximating thermodynamic quantities in the van't Hoff equation, we consider the difference in partial molar Gibbs free energy of the solute in a solid and saturated solution directly. Again, like conventional thermodynamic interpretations for solubility, the resulting model has parameters related to thermodynamic quantities other than solubility, although the solubility at a chosen reference point is required. Further, like the approach described in Chapter 3, this approach is applicable to systems where pure solute melting properties aren't available, which

is a potential limitation of conventional approaches. Finally, the approach presented in this chapter gives an analytic (but still approximate) model for solubility expressed in terms of thermodynamic quantities associated to the chosen reference point on the solubility curve, eliminating the need to perform numerical integration required to make predictions with the approach proposed in Chapter 3.

Using this approach, we estimate the temperature dependent aqueous solubility of three organic compounds: α-glycine, sucrose and urea, using a broad range of experimentally determined thermodynamic properties reported in the literature, all of which at 298.15 K. For both urea and sucrose, there exists a substantial amount of directly measured solubility data reported in the literature across a broad range of temperature when compared to α-glycine, accompanied by a meaningful amount thermodynamic data from which the quantities required to make predictions can be derived. Furthermore, not only does each system represent different chemistry, the solubility curve for each system presents a different behavior when viewed from the perspective of deviations from the van't Hoff equation. As shown, solubility predictions derived from experimental data via mathematical modeling lead to predictions in good agreement with directly measured solubility data for all compounds up to at least 340 K; however, as with the results derived for the aqueous solubility of α-glycine presented in Chapter 3, predictions at temperatures above this are sensitive to the mathematical modeling approach taken.

The remainder of this chapter is structured as follows. In the following section, a local expansion view of solubility is developed starting from the condition that the chemical potential of a solute as a pure solid and in solution are equal (i.e., Eq. (2.31)) at thermodynamic equilibrium. This includes a discussion on the local behavior of the aqueous solubility for a broad collection of organic compounds. Then, solubility estimates for each system (i.e., α-glycine, sucrose and urea) are presented in the context of available experimental data used to evaluate relevant thermodynamic quantities need to make predictions.

4.2 A Local Expansion View of Solubility

In Chapter 2, it was shown that a necessary condition for solid-liquid equilibrium was the equivalence of chemical potential between a hypothetical solid phase and a saturated liquid solution containing the same compound as that in the pure solid phase (i.e., Eq. (2.31)), which are assumed to depend on temperature, and temperature and composition, respectively. If we assume that $\beta(\mu_1^{\mathrm{l}}(T,x) - \mu_1^{\mathrm{s,\circ}}(T)) = f(T,x)$ (where subscript 1 and 2 denote solute and solvent respectively, as in Chapter 3) is adequately described (at least locally) by a Taylor series expanded around an arbitrary temperature and composition in terms of β and $\ln x$ and neglecting terms

above 1$^{\text{st}}$-order gives:

$$\begin{aligned}\beta\left(\mu_1^{l}(T,x)-\mu_1^{s,\circ}(T)\right) \approx \beta_0\left(\mu_1^{l}(T_0,x_0)-\mu_1^{s,\circ}(T_0)\right) \\ +\left(\bar{h}_1^{l}(T_0,x_0)-\bar{h}_1^{s,\circ}(T_0)\right)(\beta-\beta_0) \\ +\left(1+\frac{\partial \ln\gamma_1(T_0,x_0)}{\partial \ln x_1}\right)\left(\ln\frac{x_1}{x_{1,0}}\right),\end{aligned} \tag{4.1}$$

which, when combined with Eq. (2.31) and re-arranged, gives an explicit but approximate model describing the temperature dependence of solubility:

$$\ln\frac{x_{1,\text{sat}}(T)}{x_{1,\text{sat}}(T_0)} = -\frac{\beta_0\left(\bar{h}_1^{\text{l}}(T,x_{1,\text{sat}}(T))-\bar{h}_1^{\text{s},\circ}(T)\right)}{1+\partial \ln\gamma_1(T,x_{1,\text{sat}}(T))/\partial \ln x_1}\left(\frac{\beta}{\beta_0}-1\right) \tag{4.2}$$

where $\partial(\beta\mu_1)/\partial\beta = \bar{h}_1$ and $\partial(\beta\mu_1^{l})/\partial \ln x_1 = 1+\partial \ln\gamma_1/\partial \ln x_1$ where $\beta\mu_1^{l} = \beta\mu_1^{l,\circ} + \ln(x_1\gamma_1)$, were used. Crucially, as shown in Chapter 3, these thermodynamic quantities can be estimated via mathematical modelling of conventional experimental measurement techniques, none of which require knowledge of the pure solid melting properties, thus mitigating potential issues with predictive approaches based on Eq. (3.1). For notational convenience, Eq. (4.2) can be re-written as:

$$\ln x_{1,\text{sat}}^{\oplus}(T) \approx -\frac{c_0}{b_0}\varepsilon, \tag{4.3}$$

where $c_0 = \beta_0\Delta\bar{h}_1^{(0)}$, $b_0 = 1+\partial \ln\gamma_1^{(0)}/\partial \ln x_g$, $\varepsilon = \beta/\beta_0 - 1 = T_0/T - 1$ and $\ln x_{1,\text{sat}}^{\oplus}(T) = \ln(x_{1,\text{sat}}(T)/x_{1,\text{sat}}(T_0))$, where the superscript (0) denotes the property is evaluated at a temperature T_0 and the corresponding saturation composition $x_{1,\text{sat}}(T_0)$. As an aside, the limiting case where T_0 is taken as the pure solute normal melting temperature (i.e., T_m) reduces Eq. (4.3) to:

$$\ln x_{1,\text{sat}}^{\oplus}(T) \approx -\Delta h^{\text{fus}}(T_m)\varepsilon, \tag{4.4}$$

where Δh^{fus} is the pure solute normal enthalpy of fusion. The 1$^{\text{st}}$-order expansion (in terms of β and $\ln x_g$) (i.e., Eq. (4.3)) can be found by integrating the the exact form of the 'van't Hoff' equation applied to solid-liquid equilibrium [143]:

$$\frac{d\ln x_1^{\oplus}(T)}{d\varepsilon} = -\frac{\beta_0\Delta\bar{h}_1(T,x_{1,\text{sat}}(T))}{1+\partial \ln\gamma_1(T,x_{1,\text{sat}}(T))/\partial \ln x_1}, \tag{4.5}$$

under the assumption that the right hand side is constant, evaluated at T_0. Equation (4.3) implies that $\ln x_{1,\text{sat}}^{\oplus}(T)$ should be linear when plotted vs. $1/T$, which is sometimes referred to as a 'van't Hoff plot'. However, solubility plotted in a van't Hoff plot is generally not linear when covering a broad range of temperature [73], as alluded to in Chapter 3. To illustrate this, the aqueous solubility of thirty seven organic compounds using data reported in [6] and used previously by [144] was used. As an aside, solubility data used in this analysis is reported for all compounds at the following temperatures $T \in [273.15, 283.15, 293.15, 303.15, 313.15, 333.15, 353.15, 373.15]K$. It should be noted that the original source of the data is not specified; however, we assume it to be

reliable enough for illustrative purposes. Following the approach used in [144], solubility data for each compound was correlated using the following three parameter equation for solubility:

$$\ln x_{1,\text{sat}}(T) = A + B/T + C \ln T, \tag{4.6}$$

where A, B and C are free parameters. Once parameterised, the ratio c_0/b_0 was derived for each compound at 313.15 K via:

$$\frac{c_0}{b_0} = C - \frac{B}{T_0} \tag{4.7}$$

which comes from Eq. (4.3). Then, the reduced solubility (i.e., $\ln x^{\oplus}_{1,\text{sat}}$) was plotted against $(b_0/c_0)\varepsilon$ (where $\varepsilon = 313.15/T - 1$), fixing the reference solubility and associated slope for all compounds and providing a simple way to visualise how solubility may (or may not) deviate from a hypothetical van't Hoff behaving system as, from Eq. (4.5), a compound whose solubility followed the van't Hoff equation would lie on a diagonal line with gradient -1, and any line deviating from this would require a more complex model. The solubility data for each compound relative to it's predicted solubility via Eq. (4.3) is shown in Fig. 4.1 and a complete list of compounds is provided in Table 4.1.

From Fig. 4.1, the 1$^{\text{st}}$-order solubility model (i.e., Eq. (4.3)) has limited practical utility for many of the compounds analysed, as shown by the blue and red solid lines deviating from the dashed grey line as $c_0/b_0\varepsilon$ moves from 0 (i.e., as the solubility moves away from the chosen reference point). In the context of the Taylor series, this is generally what we expect, given that only 1$^{\text{st}}$-order terms have been used and a truncated Taylor series is approximate by definition beyond a application specific distance. More generally, fourteen of the compounds have a negative curvature (i.e., deviate 'down' from the van't Hoff reference line and colored blue) and the remaining twenty three have a positive curvature (i.e., deviate 'up' from the van't Hoff reference line and colored red). The compounds included in this analysis cover a diverse range of chemistry, including: amino acids, aromatic (alcohols, acids, etc.,), carboxylic acids, sugars, etc. For the main groups analysed, no relationship between chemistry and curvature was noted, as per Table B.1 provided in Appendix B which groups compounds by chemistry. Model parameters generated as part of this analysis are reported in Table B.2, also in Appendix B.

Given the inadequacy of Eq. (4.3), we extend our Taylor series analysis to include 2$^{\text{nd}}$-order terms for $\beta(\mu_1^l(T,x) - \mu_1^s(T))$ where the expansion is again carried out in terms of β and $\ln x_1$,

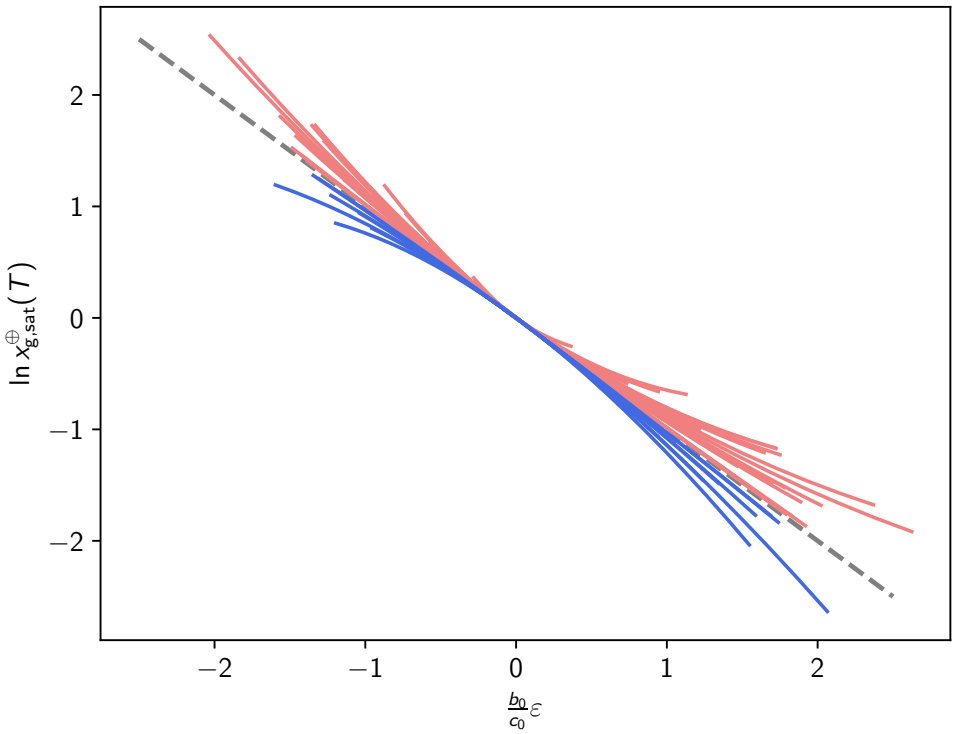


Figure 4.1: Modeled deviations from Eq. (4.3) for the aqueous solubility of 37 organic solutes reported in [6]. The dashed grey line indicates the reference for solubility obeying Eq. (4.3) at all temperatures. Solid red and blue lines correspond to modeled solubility. $\varepsilon = T_0/T - 1$, where $T_0 = 313.15\,\mathrm{K}$.

Table 4.1: Organic compounds with solubility reported in [6] and associated curvature defined relative to Eq. (4.3).

Compound	Curvature	Compound	Curvature
Adipic Acid	+	Alanine (D)	+
Alanine (DL)	+	o-Aminophenol	−
p-Aminophenol	+	Benzoic Acid	+
Citric Acid	+	Dicyandiamide	+
Fumaric Acid	+	Glucose	−
Glutamic Acid (D)	+	Glycine	−
Hydroquinone	−	o-Hydroxybenzoic Acid	+
m-Hydrobenzoic Acid	+	p-Hydroxybenzoic Acid	+
Lactose	+	Maleic Acid	−
Malic Acid (DL)	+	Malonic Acid	+
Maltose	+	Mannitol (D)	+
Melamine	+	Oxalic Acid	−
Pentaerythritol	+	o-Phthalic Acid	+
Picric Acid	+	Raffinose	−
Resorcinol	−	Salicylic Acid	+
Succinic Acid	+	Succinimide	−
Sucrose	+	Tartaric Acid	+
Tartaric Acid (racemic)	−	Taurine	−
Thiourea	−	Urea	−
Uric Acid	−		

giving:

$$\begin{aligned}\beta\left(\mu_1^l(T,x)-\mu_1^s(T)\right) \approx{}& \beta_0\left(\mu_1^l(T_0,x_0)-\mu_1^s(T_0)\right)\\ &+\beta_0\left(\bar{h}_1^l(T_0,x_0)-\bar{h}_1^s(T_0)\right)\left(\frac{\beta}{\beta_0}-1\right)\\ &-\frac{1}{2!}\frac{\left(\bar{c}_{p,1}^l(T_0,x_0)-\bar{c}_{p,1}^s(T_0)\right)}{R}\left(\frac{\beta}{\beta_0}-1\right)^2\\ &+\left(1+\frac{\partial \ln\gamma_1(T_0,x_{1,0})}{\partial \ln x_1}\right)\left(\ln\frac{x_1}{x_{1,0}}\right)\\ &+\frac{1}{2!}\left(\frac{\partial^2 \ln\gamma_1(T_0,x_{1,0})}{\partial \ln x^2}\right)\left(\ln\frac{x_1}{x_{1,0}}\right)^2\\ &+\beta_0\frac{\partial\bar{h}_1^l(T_0,x_{1,0})}{\partial \ln x_1}\left(\ln\frac{x_1}{x_{1,0}}\right)\left(\frac{\beta}{\beta_0}-1\right)\end{aligned} \tag{4.8}$$

Combining with Eq. (2.12) gives:

$$\ln\frac{x_{1,\mathrm{sat}}(T)}{x_{1,\mathrm{sat}}(T_0)} \approx \frac{-2c(T)}{b(T)+\sqrt{b(T)^2-4ac(T)}}, \tag{4.9}$$

where:

$$a=\frac{1}{2!}\frac{\partial^2 \ln\gamma_1^{(0)}}{\partial \ln x_1^2}, \tag{4.10}$$

$$b(T)=1+\frac{\partial \ln\gamma_1^{(0)}}{\partial \ln x_1}+\beta_0\frac{\partial\bar{h}_1^{l,(0)}}{\partial \ln x_1}\varepsilon=b_0+b_1\varepsilon, \tag{4.11}$$

$$c(T)=\beta_0\Delta\bar{h}_1^{(0)}\varepsilon-\frac{1}{2!}\frac{\Delta\bar{c}_{p,1}^{(0)}}{R}\varepsilon^2=c_0\varepsilon+c_1\varepsilon^2, \tag{4.12}$$

noting that $\partial^2(\beta\mu_1)/\partial\beta^2=-RT^2\bar{c}_{p,1}$ where $\bar{c}_{p,1}$ is the partial molar heat capacity of solute in solution.

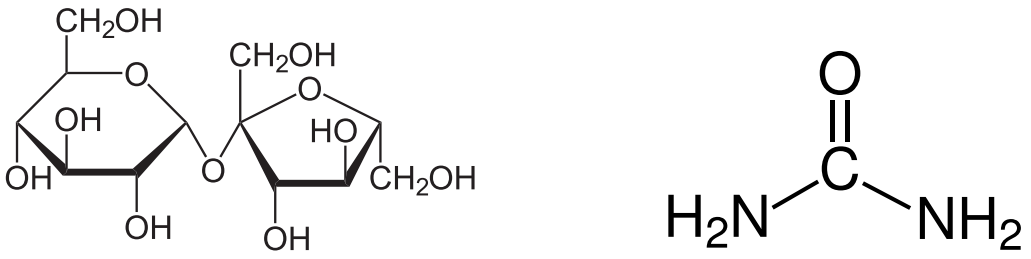
Equation (4.9) expresses solubility in terms of five thermodynamic quantities. As noted already, models with fewer parameters can correlate temperature dependent solubility data over varying temperature ranges. However, we would like to emphasize that parameters in our approach have a clear thermodynamic interpretation and can, in theory, be evaluated from experimental measurements, as illustrated in the proceeding section.

It should be noted that both Eq. (4.3) and (4.9) can be derived by considering the solute chemical potentials individually in terms of the same Taylor series expansion. However, by focusing on $\beta(\mu_1^s-\mu_1^l)$, we avoid forcing the underlying chemical potential to have specific forms, which is beyond the scope of this chapter.

4.3 The Aqueous Solubility of α-glycine, Urea and Sucrose in Water

A literature review was completed to identify model systems to test the solubility prediction framework proposed in the previous section. More specifically, systems with a significant amount of

directly measured solubility data (across a broad range of temperature), as well as solid and liquid solution phase thermodynamic data, were sought. In addition to the α-glycine-water system, which was the subject of analysis in Chapter 3, urea-water and sucrose-water were identified as model systems to examine the application of Eq. (4.9).



(a) Sucrose (b) Urea

Figure 4.2: Molecular structures of sucrose and urea.

Briefly, sucrose is a disaccharide formed from the monosaccharides fructose and glucose and it's chemical structure is illustrated in Fig. 4.2. Sucrose is an industrially relevant compound used throughout the food industry. The sucrose-water system has well defined phase boundaries and relatively speaking, when compared to other industrially relevant binary solute-solvent systems, broad coverage of experimental studies on the solution phase. The volume and diversity of data available for the sucrose-water system is captured by a substantial compilation provided in [145]. Compared to sucrose, urea is small molecule with a less complex structure (see Fig. 4.2). However, has substantial biological and industrial importance.

Directly measured aqueous solubility for α-glycine, urea and sucrose is presented in Fig. 4.3 showing that, although limited in number, these systems display diverse solubility behavior. Urea appears to have a solubility dependence adequately described by Eq. (4.3), as shown by measurement data lying on the diagonal black line. In comparison, α-glycine shows a mild negative deviation, while sucrose has a more extreme positive deviation.

4.3.1 Methodology

As with the work described in Chapter 3, there are many ways in which the thermodynamic quantities in Eq. (4.9) can be estimated from experimental measurements. In this chapter, we apply the approach used in Chapter 3 where all parameters are evaluated with thermodynamic data collected at the chosen reference temperature (i.e., T_0), which in this chapter is taken as 298.15 K (as in Chapter 3). This is in part motivated by the availability of thermodynamic data, where the majority of experimentally determined thermodynamic quantities used to make predictions are reported at standard conditions (i.e., 298.15 K and 1 atm). However, as mentioned in Chapter 3, this approach has an advantage that, if a temperature is found where all measurements are mu-

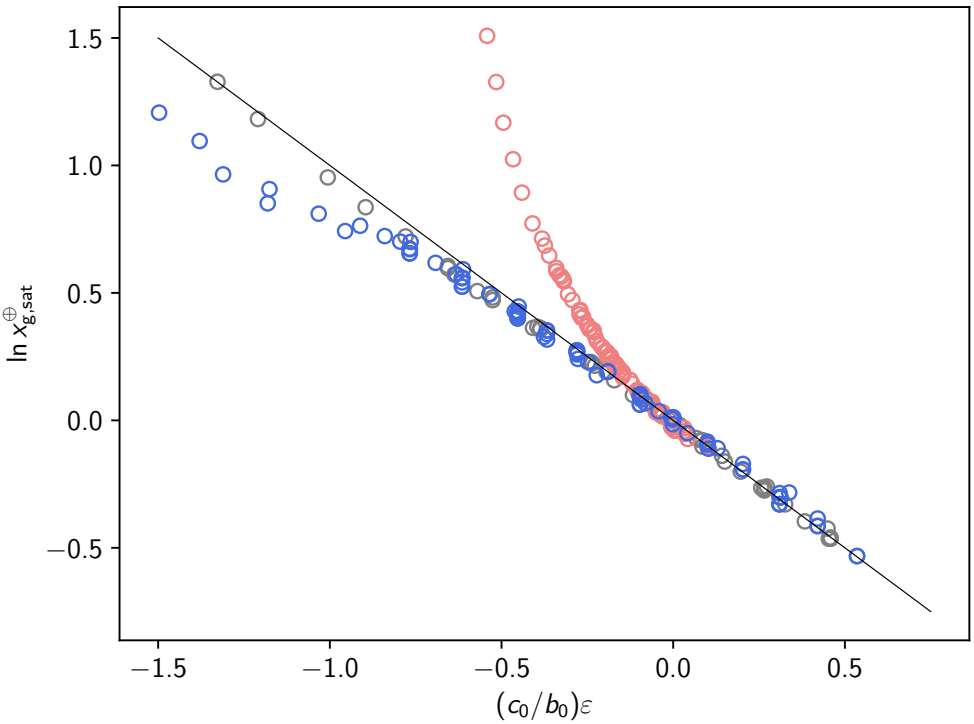


Figure 4.3: Aqueous solubility of α-glycine [89, 12, 93, 91, 92, 96, 97, 98, 99, 100, 101, 102, 25], sucrose [146, 147, 148, 149, 150, 151] and urea [152, 153, 154, 155, 156, 157, 158]. Blue open circles correspond to direct measurements for α-glycine. Grey open circles correspond to direct measurements for urea. Red open circles correspond to direct measurements for sucrose. Solid black line with slope -1 corresponds to Equation (4.3). $\ln x^{\oplus}_{g,sat} = \ln(x_{g,sat}(T)/x_{g,sat}(T_0))$ and $\varepsilon = T_0/T - 1$ where $T_0 = 298.15\,\mathrm{K}$.

tually convenient to perform (e.g., 298.15 K) and potential users of this approach have access to required instrumentation, experimental effort may be reduced. In addition, it allows estimation of quantities using measurements closely related to the property of interest. For example, we used enthalpy-based measurements to estimate the partial molar enthalpy and heat capacity measurements to estimate the partial molar heat capacity. Potential users of this approach may use any thermodynamically consistent methodology to estimate thermodynamic quantities required to make solubility predictions via Eq. (4.9).

For quantities derived from measures of solution thermodynamics, binary solutions were modeled by considering the Gibbs free energy of the solution, defined here as:

$$\beta g^{\text{sol.}} = \beta g^{\text{R}} + \beta g^{\text{E}}, \tag{4.13}$$

where βg^{R} was taken as the ideal solution defined with respect to Raoult's Law [159]:

$$\beta g^{\text{R}} = x_1(\beta \mu_1^{l,\circ} + \log x_1) + x_2(\beta \mu_2^{l,\circ} + \log x_2), \tag{4.14}$$

and βg^{E} was assumed to follow a Redlich-Kister polynomial [40]:

$$\beta g^{\text{E}} = x_1 x_2 \left[A + B(x_1 - x_2) + C(x_1 - x_2)^2 + D(x_1 - x_2)^3 + \dots\right]. \tag{4.15}$$

For a given system, it is not possible to know the most appropriate combination of parameters in Eq. (4.15) required to describe available thermodynamic data *a priori*. As such an iterative approach was employed in this chapter where the number of parameters used were increased, starting with A only, then A and B, then A, B and C, and finally A, B, C and D, and the appropriateness of each model analysed through a measure of 'best fit', as well as through visual comparison to available data, where possible.

For the purpose of this chapter, parameters in Eq. (4.15) were assumed to depend on temperature via: $M = M_0 + M_1(T_0/T - 1) + M_2(\log(T_0/T) - T_0/T + 1)$ (where M denotes A, B, C and D) with $T_0 = 298.15$ K. Interpreting the parameters in this way makes it possible to separately evaluate parameters associated different types of thermodynamic quantities required to make predictions via Eq. (4.9), in a manner similar to the approach described in Chapter 3. More specifically, for the system described by a Redlich-Kister expansion with parameters A, B, C and D, the enthalpy and heat capacity of a binary mixture at T_0 is given by:

$$\begin{aligned} h^{\text{sol.}}(T_0, x_1) = {} & x_1 h_1^{l,\circ}(T_0) + x_2 h_2^{l,\circ}(T_0) \\ & + x_1 x_2 \left[A_1 + B_1(x_1 - x_2) + C_1(x_1 - x_2)^2 + D_1(x_1 - x_2)^3\right], \end{aligned} \tag{4.16}$$

and

$$\begin{aligned} c_p^{\text{sol.}}(T_0, x_1) = {} & x_1 c_{p,1}^{l,\circ}(T_0) + x_2 c_{p,2}^{l,\circ}(T_0) \\ & + x_1 x_2 \left[A_2 + B_2(x_1 - x_2) + C_2(x_1 - x_2)^2 + D_2(x_1 - x_2)^3\right], \end{aligned} \tag{4.17}$$

respectively. For each of the systems studied in the remainder of this chapter, estimating the parameters required to make solubility predictions with Eq. (4.9) involved three stages of modeling.

First, data from which the solvent activity (defined relative to Raoult's Law) in solution could be derived was used to evaluate parameters A_0, B_0, C_0 and D_0 in Eq. (4.15), noting that

$$\left(\frac{\partial \beta G}{\partial n_2}\right)_{n_1,T,P} = \ln \gamma_2, \tag{4.18}$$

where $\ln \gamma_2$ is the solvent activity coefficient (again, defined relative to Raoult's Law). Then, estimates for the solute activity coefficient in solution at discrete concentrations were determined using the parameters just mentioned via:

$$\left(\frac{\partial \beta G}{\partial n_1}\right)_{n_j,T,P} = \ln \gamma_1. \tag{4.19}$$

Both steps were performed using a bespoke Python code leveraging Automatic Differentiation, described in greater detail in Chapter 5. To estimate the composition partial derivatives required to make predictions (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$), numerical differentiation was used, noting:

$$\frac{\partial f(x)}{\partial x} \approx \frac{f(x+\Delta x) - f(x-\Delta x)}{2(\Delta x)}, \tag{4.20}$$

and:

$$\frac{\partial^2 f(x)}{\partial x} \approx \frac{f(x-\Delta x) - 2f(x) + f(x+\Delta x)}{(\Delta x)^2}. \tag{4.21}$$

It should be noted that a value of $\Delta \ln x_1 = 0.001$ was taken for all systems.

Data related to enthalpy changes involving the liquid solution phase were used to parameterise Eq. (4.16), providing estimates for A_1, B_1, C_1 and D_1. As with the parameters related to the solute activity coefficient (and associated composition derivatives), the excess partial molar enthalpy (defined relative to the infinitely dilute reference state) required to estimate $\Delta \bar{h}_1^{(0)}$ was found using AD, since:

$$\left(\frac{\partial H}{\partial n_1}\right)_{n_j,T,P} = \bar{h}_1. \tag{4.22}$$

However, as with the solute activity partial derivatives, the composition derivative (i.e., $\beta_0 \partial \bar{h_1}^{I,(0)} / \partial \ln x_1$) was found using numerical differentiation Eq. (4.20).

Finally, solution heat capacity data was used to parameterise Eq. (4.17), giving estimates for A_2, B_2, C_2 and D_2. Again, the partial molar enthalpy of solute in solution was determined using AD, given that:

$$\left(\frac{\partial C_p}{\partial n_1}\right) = \bar{c}_{p,1}. \tag{4.23}$$

For the models proposed in this chapter, all fitting was completed using a 'residual sum of squares' (RSS) objective function (i.e., $RSS = \sum_i^N (y_i^{\text{data}} - f(\theta, x_i)^{\text{model}})^2$, where y_i^{data} is an experimentally measured quantity, $f(\theta, x_i)^{\text{model}}$ is the same quantity predicted from the relevant model and N is the number of data points). It should be noted that, at each stage of the modelling

process (described above), experimental data was weighted equally (c.f. Chapter 5), with the exception for obvious outliers which were discarded when relevant. All fitting was completed using a bespoke set of Python modules (noted previously) described in Chapter 5 coupled with the Nelder-Mead optimisation algorithm [160] implemented in SciPy [161]. For parameter estimation, the optimisation algorithm used requires initial estimates for parameter values. For most parameter sets evaluated in the remainder of this chapter, satisfactory parameter values were found by initialising the optimisation algorithm with all parameter values set to zero; however, in some cases, paramater values were initialised at non-zero values to obtain reasonable fits. Generally speaking, this is to be expected on occasion and can be related to the optimisation process (i.e., algorithm, measure of best fit, etc.,) finding local minima (and failing to find 'better' solutions), subject to poor initial parameter choice. For completeness, the initial values associated to the parameter values tabulated in the remainder of this chapter are provided in Appendix C.

The remainder of this chapter is dedicated to the mathematical modelling of available experimental data, estimation of thermodynamic quantities required to make use of Eq. (4.9), and assessment of associated solubility predictions by comparison to available experimental data. It should be noted that for each system (i.e., α-glycine-water, sucrose-water and urea-water), solubility predictions are presented for all models which provide a 'reasonable' fit to experimental data. We note that from a certain perspective, this means we include models that are unnecessarily complex (i.e., additional parameters don't make a meaningful difference in terms of reducing measure of best fit); however, including a range of models provides a simple way to capture model-based variance and it's impact on the applicability of Eq. 4.9 as a predictive tool, and aides assessment of sensitivity highlighted in Chapter 3.

4.3.2 α-Glycine

For the α-glycine-water system, data from which the water activity coefficient could be inferred included six independent reports of glycine isopiestic molality as noted in Chapter 3. As before, isopiestic molalities were converted to osmotic coefficients using a model for the omsotic coefficient of reference solutions and then to solvent activity coefficient via Eq. (3.39). Experimental data and corresponding best fits derived from Eq. (4.15) with parameters A_0, B_0, C_0 and D_0 are presented in Fig. 4.4. The resulting parameter values for each fit are provided in Table 4.2.

From Fig. 4.4, the fit with only one parameter (i.e., A_0) produces a model in poor visual agreement with available data. However, models with at least two parameters (i.e., A_0, B_0, etc.,) gave fits in good visual agreement with available data. These findings are supported by the associated RSS values provided in Table 4.2, where models 2, 3 and 4 have a markedly lower RSS, when compared to model 1. Close inspection of Fig. 4.4 indicates that near α-glycine saturation (i.e., 0.056 mole fraction), the fits including more than one parameter deviate slightly, which may

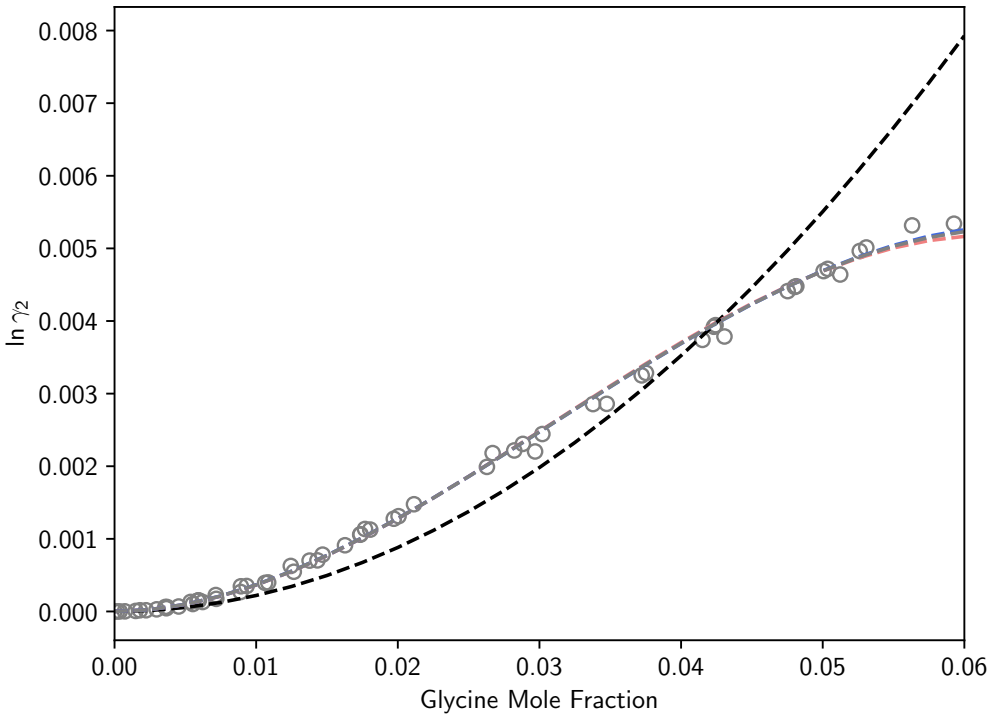


Figure 4.4: Processed water activity coefficient (Raoult's Law) data in binary mixtures with glycine at 298.15 K. Open grey symbols denote processed experimental data [107, 108, 109, 87, 110, 111, 85, 112, 113, 114, 115, 116], . Colored dashed lines correspond to fits using Eq. (4.15), with black denoting model 1, red denoting model 2, blue denoting a fit with model 3 and grey denoting model 4, as per Table 4.2.

Table 4.2: Redlich-Kister parameter values for α-glycine solvent activity data fits. Values reported for $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$ have been evaluated at solution composition of 0.056 glycine mole fraction.

Model	RSS	A_0 (-)	B_0 (-)	C_0 (-)	D_0 (-)	$\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ (-)	$\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$ (-)
1	0.0096	2.20				0.0096	0.0276
2	0.00025	-28.95	11.01			0.0024	-0.0185
3	0.00023	4.49	-8.87	5.25		0.0029	-0.0148
4	0.00023	-5.98	-2.34	3.22	0.146	0.0028	-0.0158

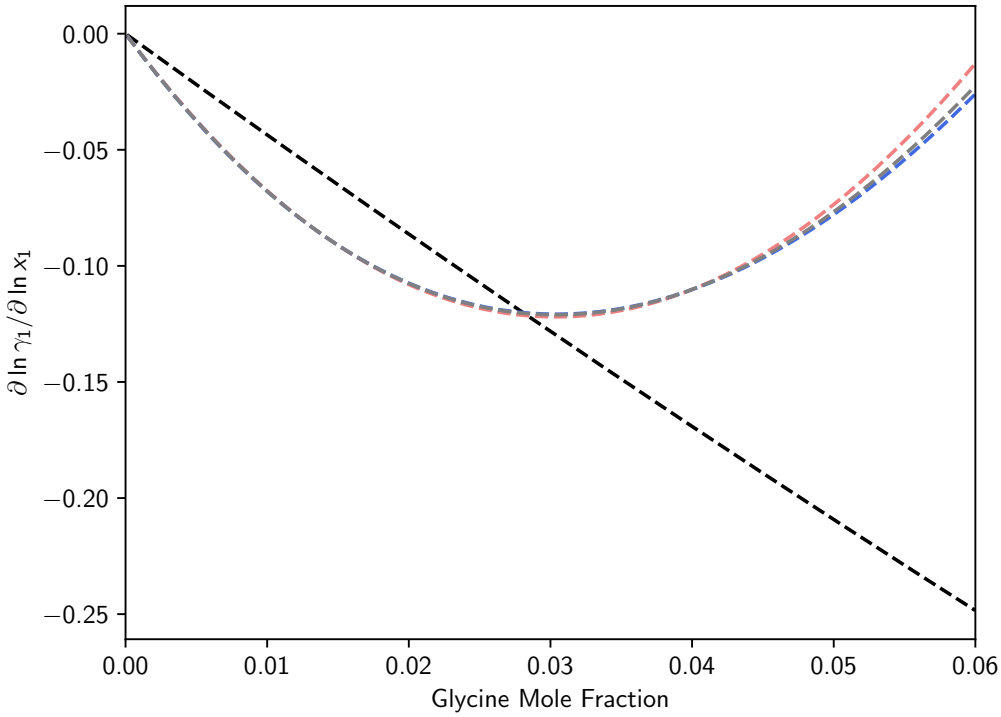


Figure 4.5: Solute activity coefficient first partial derivative with respect to solute composition estimated from best fits shown in Fig. 4.4. Colored dashed lines correspond to different models, with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

lead to differences in the partial derivatives (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) required for predictions. As illustrated in Chapter 3, differences in these quantities can lead to significant differences in predictions far from the chosen reference temperature.

The first and second solute activity coefficient partial derivatives with respect to composition derived from the fits to solvent activity data are presented in Fig. 4.5 and 4.6, respectively. The partial derivatives (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) derived from models including more than one parameter are in good agreement; however, begin to diverge close to the α-glycine saturation composition. For each model, values for $\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ and $\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$ at an assumed saturation composition of 0.056 glycine mole fraction are presented in Table 4.2.

Again, as outlined in Chapter 3, the enthalpy of dilution has been reported at 298.15 K over a range of solution compositions by at least four independent studies (based on the dataset used in [80]), all of which report dilution measurements where a finite amount of solvent was added to a glycine-water solution of known concentration and the heat evolved measured, as well as the final solution composition. Equation (4.15) was fit to the composite dataset via Eq. (3.19) and

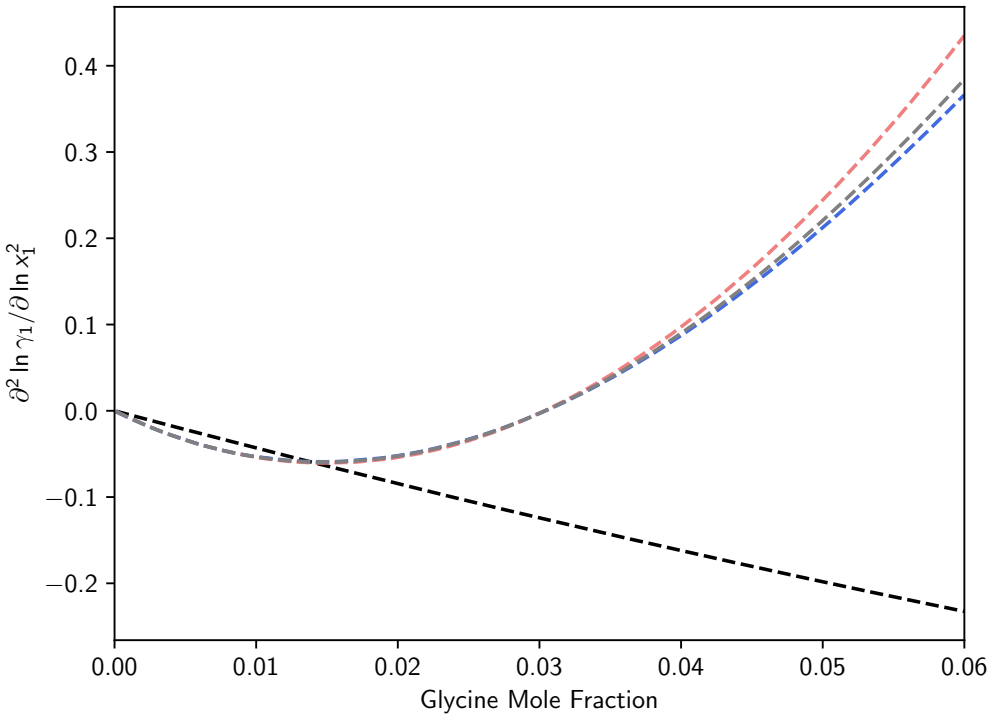


Figure 4.6: Solute activity coefficient first partial derivative with respect to solute composition estimated from best fits shown in Fig. 4.4. Colored dashed lines correspond to different models, with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

the resulting parameters are given in Table 4.3, with estimates for $\Delta\bar{h}^{(0)}$ (noting the enthalpy of solution to infinite dilution for α-glycine was taken as 14523 J/mol.K, as noted in Chapter 3) and $\beta_0\partial\bar{h}_1^{l,(0)}/\partial\ln x_1$. For each model, the corresponding estimate of $\bar{h}_1^{\mathrm{l}} - h_1^{\infty}$ is presented in

Table 4.3: Redlich-Kister parameter values for α-glycine enthalpy of dilution data fits

Model	RSS	A_1 (J/mol)	B_1 (J/mol)	C_1 (J/mol)	D_1 (J/mol)	$\Delta\bar{h}_1^{(0)}$ (J/mol)	$\beta_0\partial\Delta\bar{h}_1^{(0)}/\partial\ln x_1$ (-)
1	445151	4.66				13265	-0.49
2	409354	-59.02	21.71			13878	0.13
3	330216	8132.34	-4774.43			11732	-4.39
4	330167	277.63	1105.87	-1521.98	575.11	11789	-4.20

Fig. 4.7. The quantity $\beta_0\partial\Delta\bar{h}_1^{(0)}/\partial\ln x_1$ derived from the aforementioned fits is presented in Fig. 4.8. Focusing on the RSS values for models 1–4 in Table 4.3, as the number of parameters increases the RSS decreases with each additional parameter—potentially indicating that model 4 best describes the data. Based on each models predicted composition dependence for $\bar{h}_1^{l} - h_1^{l,\infty}$ and $\beta_o\partial\bar{h}_1^{l}\partial\ln x_1$ illustrated in Fig. 4.7 and 4.8, there are three different behaviors predicted by model 1, model 2 and models 3 and 4 (which are in good agreement with each other), respectively. Briefly, the motivation for illustrating $\bar{h}_1^{\mathrm{l}} - h_1^{\infty}$ comes from, as described previously in Chapter 3, the fact that it can be combined with enthalpy of solution measurements reported for infinitely dilute solution concentrations giving the quantity $\Delta\bar{h}^{(0)}$ which is required to make predictions. However, as noted in Chapter 3, there is no simple way to visually compare models to enthalpy of dilution data, as such it is challenging to validate whether the reduced RSS observed for models 3–4 is the result of over-fitting—where a mathematical model is too flexible for a given dataset, and captures noise rather than signal.

For $\Delta\bar{c}_{p,1}^{(0)}$, two quantities are required—the pure solid solute molar heat capacity (i.e., $\bar{c}_{p,1}^{s}(T_0)$) and the partial molar heat capacity of solute in solution at saturation (i.e., $\bar{c}_{p,1}^{l}(T_0, x_{1,\mathrm{sat}}(T_0))$). It is possible to measure the pure solid solute heat capacity using conventional techniques, and taken as 99.23 J/mol.K for α-glycine at 298.15 K (see Chapter 3). There are no methods available to measure the solute partial molar heat capacity in solution directly; however, it is possible to estimate if the solution heat capacity is available at the reference temperature as a function of solution concentration (which is readily measurable), since $\bar{c}_{p,1}^{l} = \partial C_p^{\mathrm{soln}}/\partial n_1$.

Heat capacity data for glycine-water solutions reported in the literature at 298.15 K is shown in Fig. 4.9 with best fits derived from Eq. (4.17) where the molar heat capacity of pure liquid water (i.e., $c_{p,2}^{l,\circ}$) was taken as 75.2 J/mol.K. It should be noted that fitting solution heat capacity data also requires fitting an additional parameter—$c_{\mathrm{p,1}}^{\mathrm{l},\circ}(T_0)$ (as well as of A_2, B_2, etc.)—which

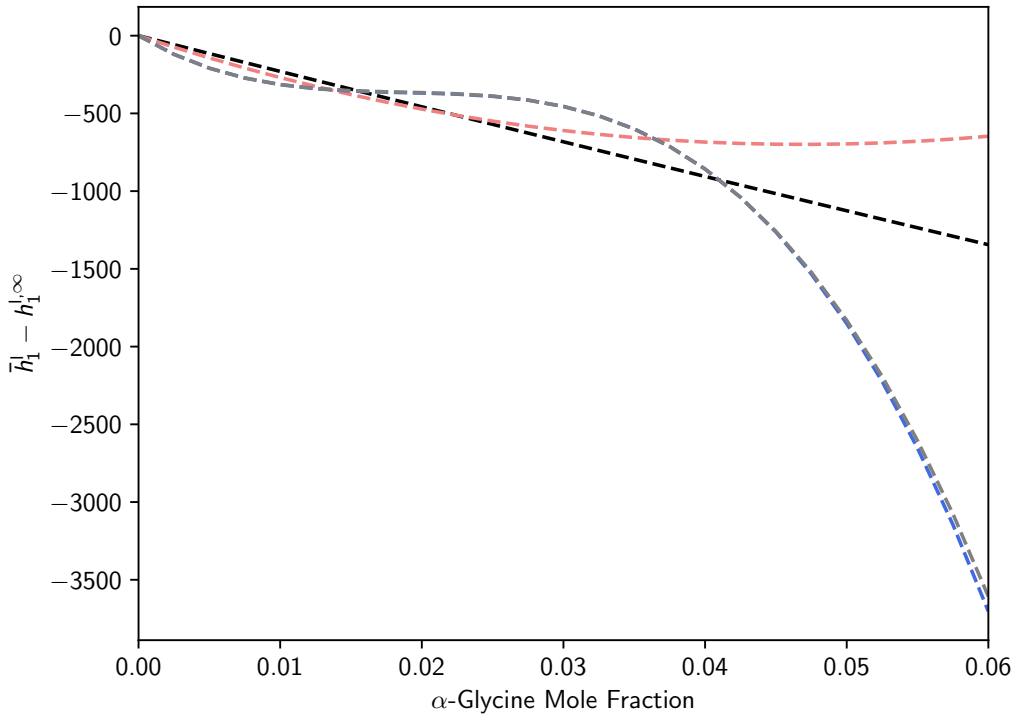


Figure 4.7: Estimates for $\bar{h}_1^{\mathrm{l}} - h_1^{\mathrm{l},\infty}$ derived from fits to experimental enthalpy of dilution measurements at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

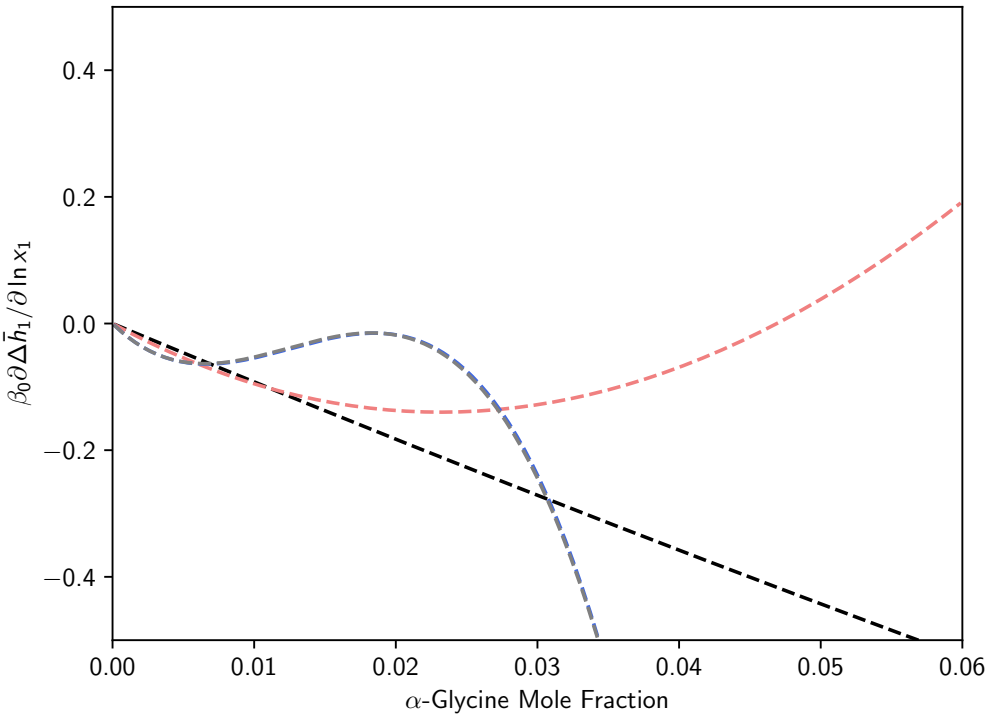


Figure 4.8: Estimates for $\beta\partial\Delta\bar{h}_1^{\mathrm{l}}/\partial\ln x_1$ derived from fits to experimental enthalpy of dilution measurements at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

may be interpreted as the heat capacity of pure liquid solute.

Experimentally determined solution heat capacity measurements for glycine-water solutions are shown alongside best-fit models in Fig. 4.9. The resulting parameter values for each fit are

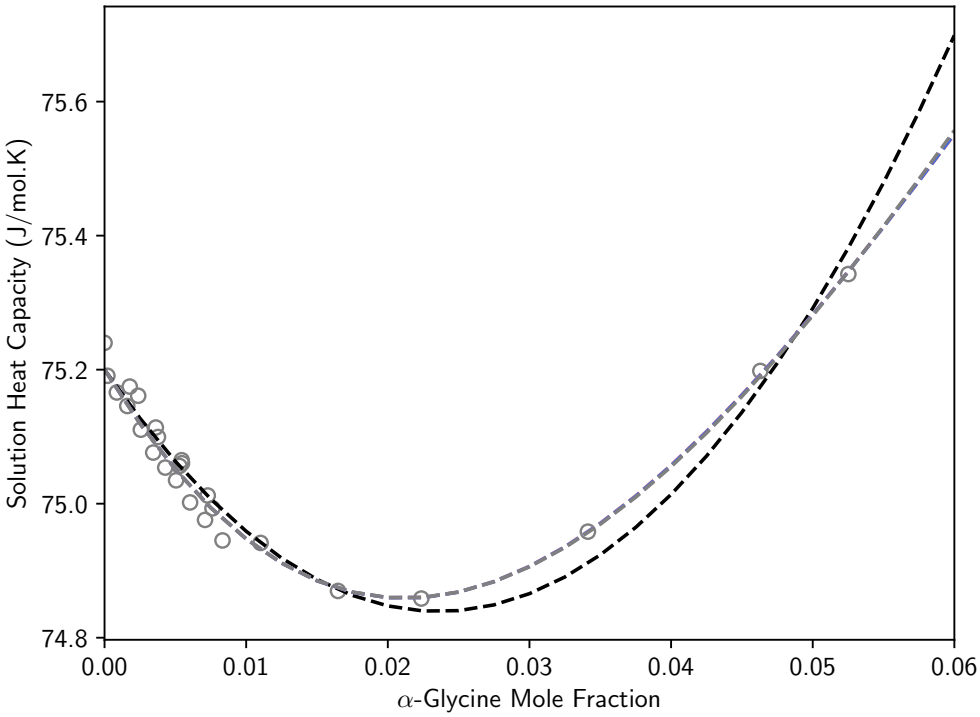


Figure 4.9: Solution heat capacity for glycine-water solutions at 298.15 K. Open grey circles correspond to solution heat capacity measurements reported in literature [77, 125, 126, 127, 128]. Colored dashed lines correspond to different models, with black denoting a fit with A_2, red denoting a fit with A_2 and B_2, blue denoting a fit with A_2, B_2 and C_2 and grey denoting a fit with A_2, B_2, C_2 and D_2. It should be noted that red, blue and grey lines overlap.

provided in Table 4.4, with estimates for $\Delta\bar{c}_{p,1}^{(0)}$.

Table 4.4: Redlich-Kister parameter values and for glycine-water solution heat capacity fits

Model	RSS	A_2 (J/mol.K)	B_2 (J/mol.K)	C_2 (J/mol.K)	D_2 (J/mol.K)	$c_{p,1}^{s,\circ}$ (J/mol.K)	$\Delta\bar{c}_{p,1}^{(0)}$ (J/mol.K)
1	0.016	-648.61				693.21	15.98
2	0.010	6589.24	-2521.38			-4027.16	1.93
3	0.010	3125.67	-476.47	-535.18		-2073.35	2.12
4	0.010	519.75	-432.44	925.92	-692.68	-279.98	2.74

From Fig. 4.9, all models are in good visual agreement with available solution heat capacity data. It should be noted that models 2–4 give models in excellent agreement with each other, but noticeably different to model 1. With reference to the RSS values given in Table 4.4, models 2–4 have a lower value compared to model 1. The partial molar heat capacity of glycine derived from each model is presented in Fig. 4.10, where it is shown that the models able to represent available data give slightly different estimates at α-glycine saturation (i.e., 0.056 mole fraction). As shown, the partial molar heat capacity of sucrose is in excellent agreement for models 2–4, but

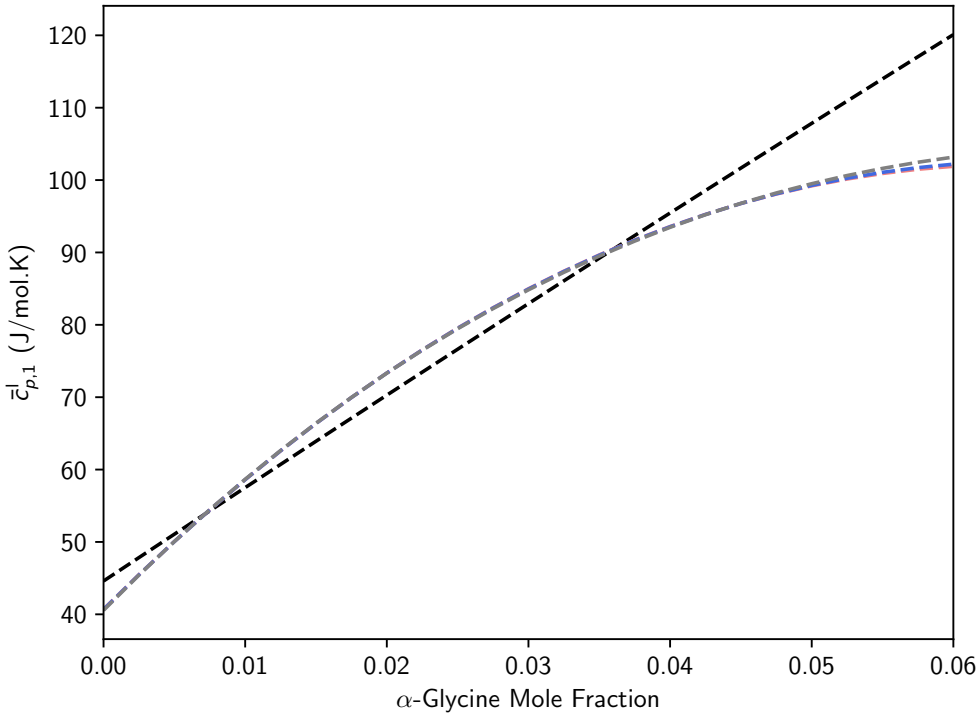


Figure 4.10: Partial molar heat capacity of glycine in glycine-water mixtures at 298.15 K derived from fits to soution heat capacity data using Redlich-Kister models. Colored dashed lines correspond to different models, with black denoting a fit with A_2, red denoting a fit with A_2 and B_2, blue denoting a fit with A_2, B_2 and C_2 and grey denoting A_2, B_2, C_2 and D_2. It should be noted that red, blue and grey lines overlap.

differs for model 1.

Thus far, different types of experimental data have been used to develop estimates for the thermodynamic quantities required to estimate solubility via Eq. (4.9). As shown, the modeling process used leads to different estimates for thermodynamic quantities, even when different models appear to have nearly equivalent fits to available data, from a RSS and visual perspective. For parameters related to the solute activity coefficient (i.e., $\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) and

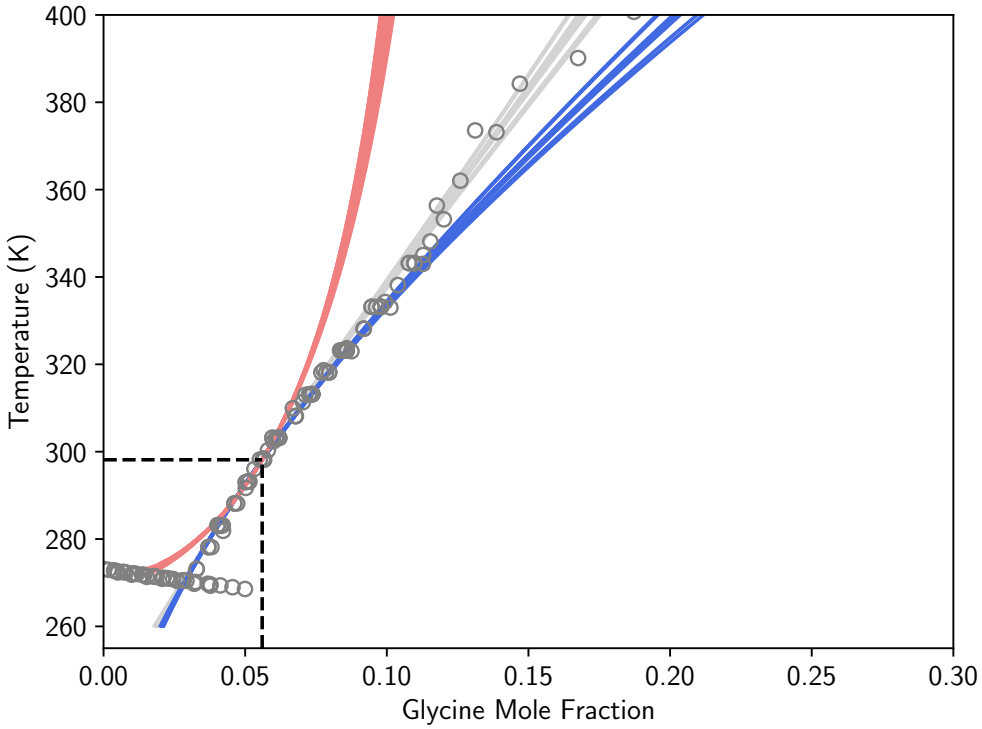


Figure 4.11: α-Glycine solubility predictions based on Eq. (4.9). Open grey symbols denote directly measured solid-liquid equilibria reported in the literature (solubility [89, 12, 93, 91, 92, 96, 97, 98, 99, 100, 101, 102, 25] and freezing point [131, 132, 133, 134, 135]). Solid lines are predictions derived from Eq. (4.9) based on all combinations of thermodynamic quantities provided in Table 4.2 (models 2–4), 4.3 (all models) and 4.4 (all models). The red band corresponds to all predictions based on model three and four in Table 4.3. The grey band corresponds to all predictions based on model one in Table 4.3. The blue band corresponds to all predictions based on model two in Table 4.3.

partial molar heat capacity (i.e., $\Delta\bar{c}_{p,1}^{(0)}$), available experimental data provides a degree of certainty around their value, despite there being differences in derived quantities between models providing 'good' fits to experimental data. In contrast, there is significant uncertainty with parameters derived from enthalpy measurements (i.e., $\Delta\bar{h}_1^{(0)}$ and $\beta_0\partial\Delta\bar{h}_1^0/\partial\ln x_1$), due to the availability and type of experimental data, and our inability to visually assess fits.

Solubility predictions based on Eq. (4.9) are presented in Fig. 4.11. The predictions correspond to all combinations of derived quantities based on models that gave plausible fits to experimental data, including: models 2–4 in Table 4.2, all models in Table 4.3 and all models in Table 4.4

As shown, predictions are grouped into three distinct 'bands' driven by the values of $\Delta\bar{h}_1^{(0)}$ and $\beta_0\partial\Delta\bar{h}_1^{(0)}/\partial\ln x_1$. Predictions in the red band are based on enthalpy parameters derived from

model three and four in Table 4.3 and are in extremely poor agreement with available solubility data, even locally around 298.15 K. This potentially indicates that these models were too flexible for the available data, leading to over-fitting.

Predictions in the grey band are based on enthalpy parameters derived from model one and predictions in the blue band are based on enthalpy parameters derived from model two in Table 4.3, respectively. Both bands give predictions in good agreement with directly measured solubility data near the reference temperature (i.e., in the range 270–340 K, where data is most reliable). Interestingly, across nearly the entire range of temperature for which directly-measured solubility data is available (i.e., 270–390 K), models in the grey band are in excellent agreement with said data.

Beyond 340 K, predictions in the grey and blue bands diverge in a similar manner to those presented in Chapter 3. However, beyond the differences in the underlying mathematical model used for fitting, the thermodynamic quantities driving divergence are different (i.e., in Chapter 3, it was quantities related to the composition dependence of the solute activity coefficient, while here it is the enthalpy). Closer inspection of the values of $\Delta\bar{h}_1^{(0)}$ and $\beta_0\partial\Delta\bar{h}_1^{(0)}/\partial \ln x_1$ in Table 4.3 suggest that it's the latter driving the divergence between grey and blue bands, given the significant difference depending on the chosen model.

In summary, solubility predictions based on Eq. (4.9), where all thermodynamic properties required were derived from experimentally measurable quantities reported at 298.15 K, give encouraging results for the aqueous solubility of α-glycine. However, additional measurements from which $\Delta\bar{h}_1^{(0)}$ and $\beta_0\partial\bar{h}_1^{l}/\partial \ln x_1$ can be derived would be useful to reduce divergence between predictions at temperatures above 340 K, and disentangle the effect of uncertainty due to the approximate nature of Eq. (4.9), and that due to mathematical modelling as a tool to estimate required quantities.

4.3.3 Sucrose

For the sucrose-water system, the types of experimental measurements reported include: vapor pressure, isopiestic molality (i.e., osmotic coefficients) and relative humidity [145].

For completeness, vapor pressure measurements of pure water above sucrose-water solutions, the water activity coefficient was found using via Eq. (3.41). When a dataset didn't include a reference vapor pressure measurement for pure water, the value 3169 Pa was used. For measurements of the isopiestic molality, the water activity coefficient was found in the same way as described for α-glycine solutions. For measurements of the relative humidity, the water activity coefficient was found via:

$$\ln \gamma_2 = \ln\left(\frac{ERH}{100 x_2}\right). \tag{4.24}$$

Experimental data processed into $\ln \gamma_w$, and corresponding best fits derived from Eq. (4.15) with parameters A_0, B_0, C_0 and D_0 are presented in Fig. 4.12. Parameter values for each fit are

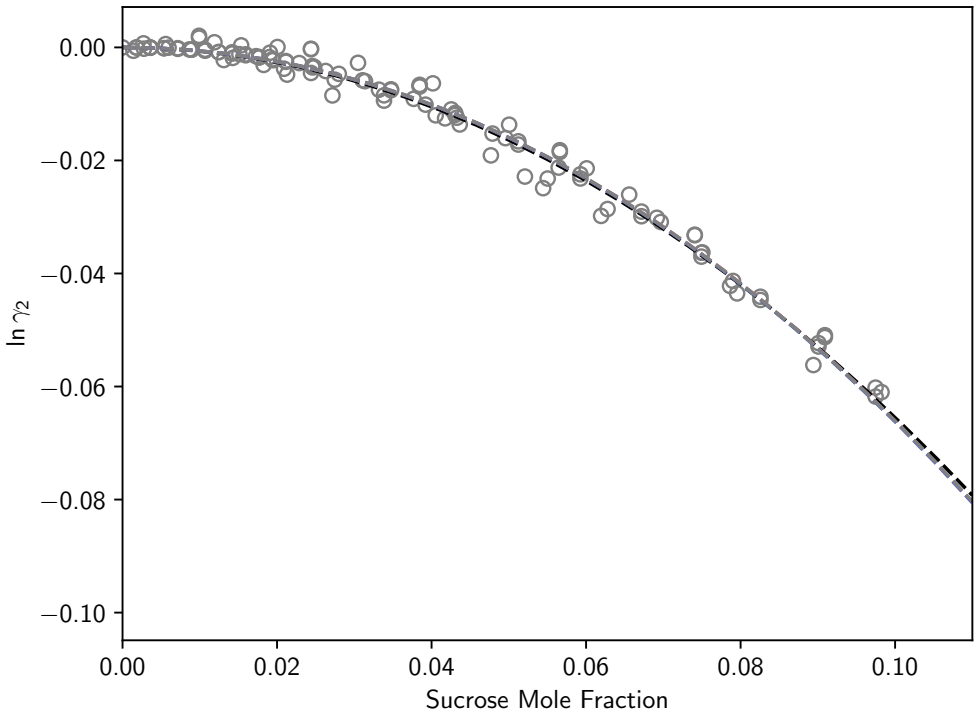


Figure 4.12: Processed water activity coefficient (Raoult's Law) data in sucrose-water liquid mixtures at 298.15 K. Open grey symbols denote estimates from processed experimental data [162, 163, 164, 165, 166, 167, 168, 169, 170, 150, 171, 172, 173, 174]. Colored dashed lines correspond to fits using Eq. (4.15), with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

provided in Table 4.5. As illustrated in Fig. 4.12, all fits produce models in good visual agreement

Table 4.5: Redlich-Kister parameter values for sucrose solvent activity data fits

Model	RSS	A_0 (-)	B_0 (-)	C_0 (-)	D_0 (-)	$\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ (-)	$\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$ (-)
1	0.0004	-6.55				1.158	1.032
2	0.0004	-9.55	1.13			1.207	1.193
3	0.0004	-2.13	-3.5	1.31		1.202	1.146
4	0.0004	0.44	-5.08	1.74	0.01	1.199	1.127

with available data. Close inspection of Fig. 4.12 indicates that near sucrose saturation (i.e.,

0.098 mole fraction), the fits including more than one parameter deviate slightly, which may lead to differences in the partial derivatives (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) required for predictions.

The first and second solute activity coefficient partial derivatives with respect to composition (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) derived from the fits to solvent activity data are presented in Fig. 4.13 and 4.14, respectively. From Fig. 4.13, the first partial derivate estimates for all

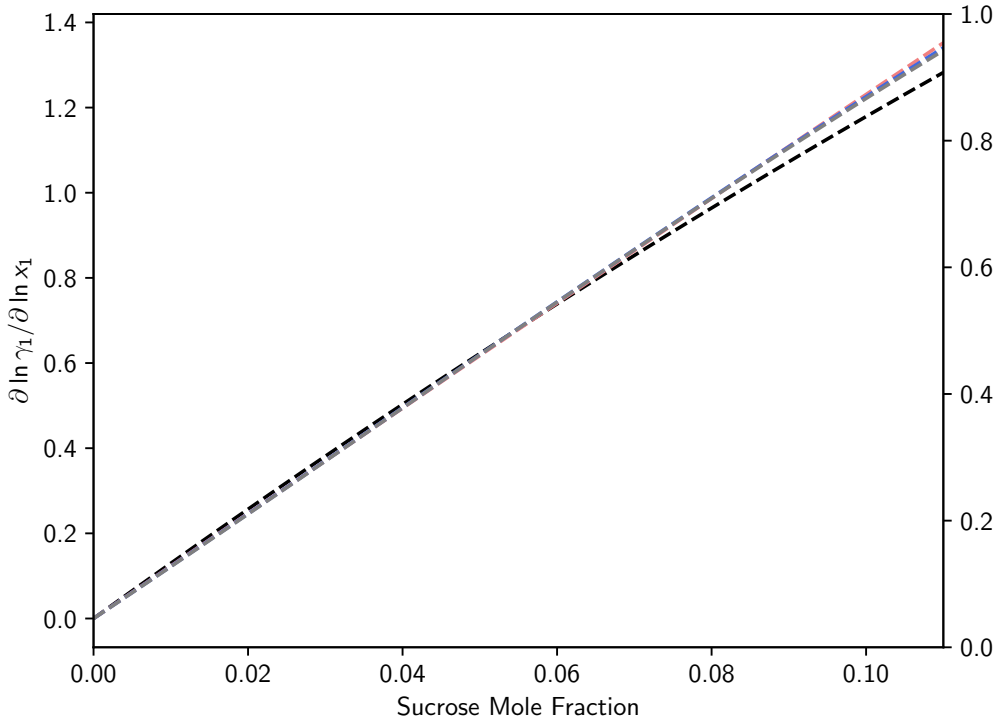


Figure 4.13: Solute activity coefficient first partial derivative with respect to solute composition estimated from best fits shown in Fig. 4.12. Colored dashed lines correspond to different models, with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

models are in good agreement; however, give different estimates at sucrose saturation. From Fig. 4.14, the second partial derivative estimates for all models are in good agreement at low compositions, but begin to diverge at approximately 0.04 sucrose mole fraction, leading to different estimates at saturation.

For sucrose, as with α-glycine, parameters in Eq. (4.15) are evaluated by correlating different types of enthalpy of dilution data reported in the literature at 298.15 K and over a range of solution concentrations and combining with reported values for the enthalpy of solution to infinite dilution, in the absence of enthalpy of solution measurements at higher solution compositions. Strictly

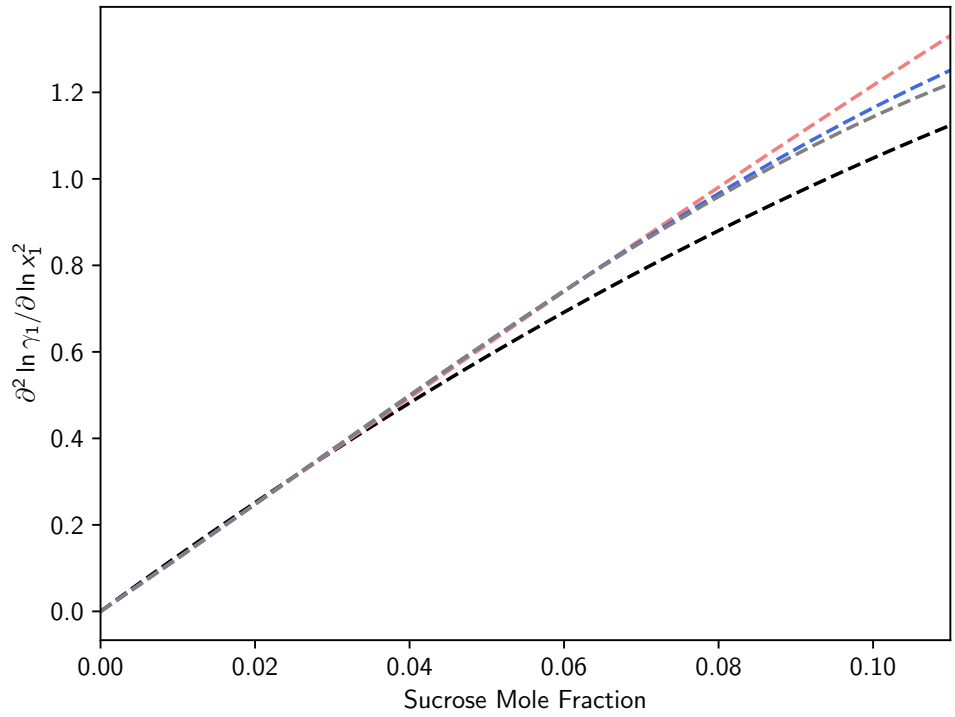


Figure 4.14: Solute activity coefficient first partial derivative with respect to solute composition estimated from best fits shown in Fig. 4.12. Colored dashed lines correspond to different models, with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

speaking, only one report of enthalpy of dilution data at 298.15 K was found in the literature [175], covering dilute solutions (i.e., up to approximately 0.04 sucrose mole fraction). As an aside, this data reports the heat evolved when two sucrose-water solutions of known composition were mixed, contrary to the dilution measurements discussed for α-glycine, where pure water was added to a solution of known composition. As such, parameters were found via:

$$\Delta H^{\text{dil}} = \left(\sum_i^N n_i \bar{h}_g^l\right)^{\text{final}} - \left[\left(\sum_i^N n_i \bar{h}_g^l\right)^1 + \left(\sum_i^N n_i \bar{h}_g^l\right)^2\right] \tag{4.25}$$

$$= \left(\sum_i^N n_i \frac{\partial \ln \gamma_i}{\partial \beta}\right)^{\text{final}} - \left[\left(\sum_i^N n_i \frac{\partial \ln \gamma_i}{\partial \beta}\right)^1 + \left(\sum_i^N n_i \frac{\partial \ln \gamma_i}{\partial \beta}\right)^2\right], \tag{4.26}$$

To supplement data from [175], data reported at 299.15 K in [176] was used. It should be noted that this data was presented in the same form as that for α-glycine, and processed in the same manner. The resulting parameters derived from fitting enthalpy of dilution data are given in Table 4.6. For each model, the corresponding estimate of $\bar{h}_1^{\text{l}} - h_1^{\text{l},\infty}$ is presented in Fig. 4.15,

Table 4.6: Redlich-Kister parameter values for sucrose enthalpy of dilution data fits

Model	RSS	A_1 (J/mol)	B_1 (J/mol)	C_1 (J/mol)	D_1 (J/mol)	$\Delta\bar{h}_1^{(0)}$ (J/mol)	$\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$ (-)
1	4090.56	-11.44				11285	2.02
2	2018.95	18.75	-10.45			10427	1.14
3	2002.8	-95.85	57.75	-18.05		10569	1.51
4	2001.3	7.29	-20.92	20.04	-8.23	10568	1.49

and the quantity $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$ derived from the aforementioned fits is presented in Fig. 4.16. From Fig. 4.15 and 4.16, estimates of $\bar{h}_1^{\text{l}} - h_1^{\text{l},\infty}$ and $\beta_0 \partial \Delta \bar{h}_1^{l} / \partial \ln x_1$ are partitioned into three groups depending on the number of parameters used. More specifically, both model one and two have distinct behaviors, whereas models 3 and 4 have very similar behaviour but distinct to models 1 and 2. It should be noted that, although models three and four give consistent estimates for $\bar{h}_1^{\text{l}} - h_1^{\text{l},\infty}$, estimates for $\beta_0 \partial \Delta \bar{h}_1 / \partial \ln x_1$ diverge very slightly near sucrose saturation (i.e., 0.098 mole fraction). Values of $\Delta \bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$ required to make predictions using Eq. (4.9) derived from each model are given in Table 4.6. It should be noted that estimates of $\Delta \bar{h}^{(0)}$ were calculated in a similar manner as described for α-glycine, using the enthalpy of solution to infinite dilution for sucrose taken as 6000 J/mol.

Additional estimates for $\bar{h}_1^{l} - \bar{h}_1^{l,\infty}$ and $\beta_0 \partial \Delta \bar{h} / \partial \ln x_1$ were derived from a comprehensive modeling study for the sucrose-water system reported by [145]. In this work, a model for the sucrose-water solution was developed by simultaneously fitting all available thermodynamic quantities,

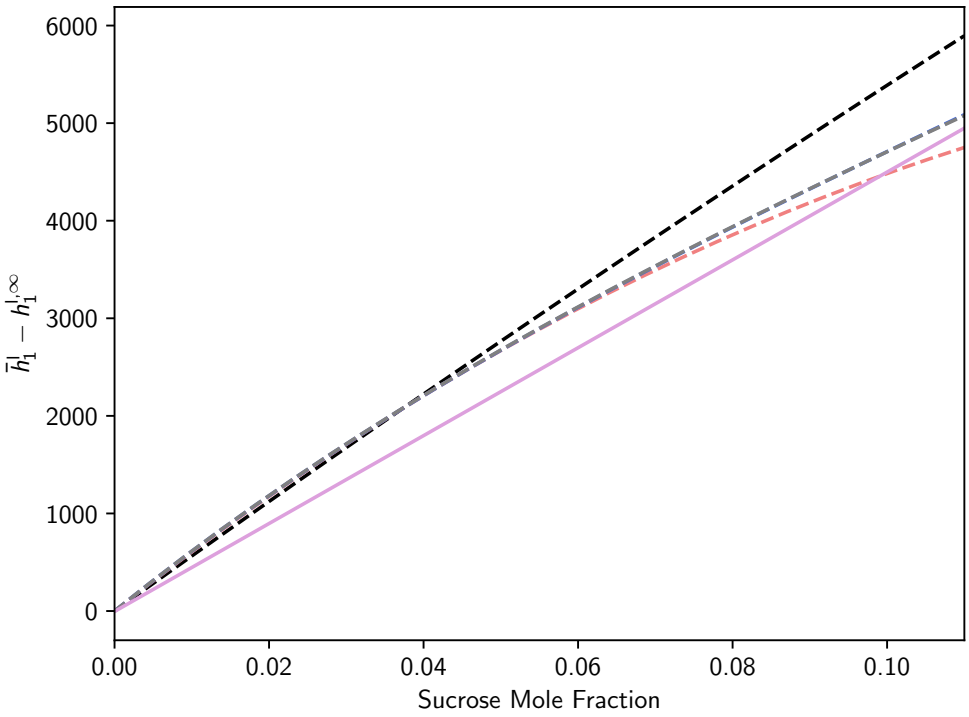


Figure 4.15: Estimates for $\bar{h}_1^{\mathrm{l}} - h_1^{\mathrm{l},\infty}$ derived from fits to experimental enthalpy of dilution measurements for sucrose-water mixtures at 298.15 K [175] and 299.15 K [176] using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1, grey denoting a fit with A_1, B_1, C_1 and D_1, and purple denoting estimates from [145].

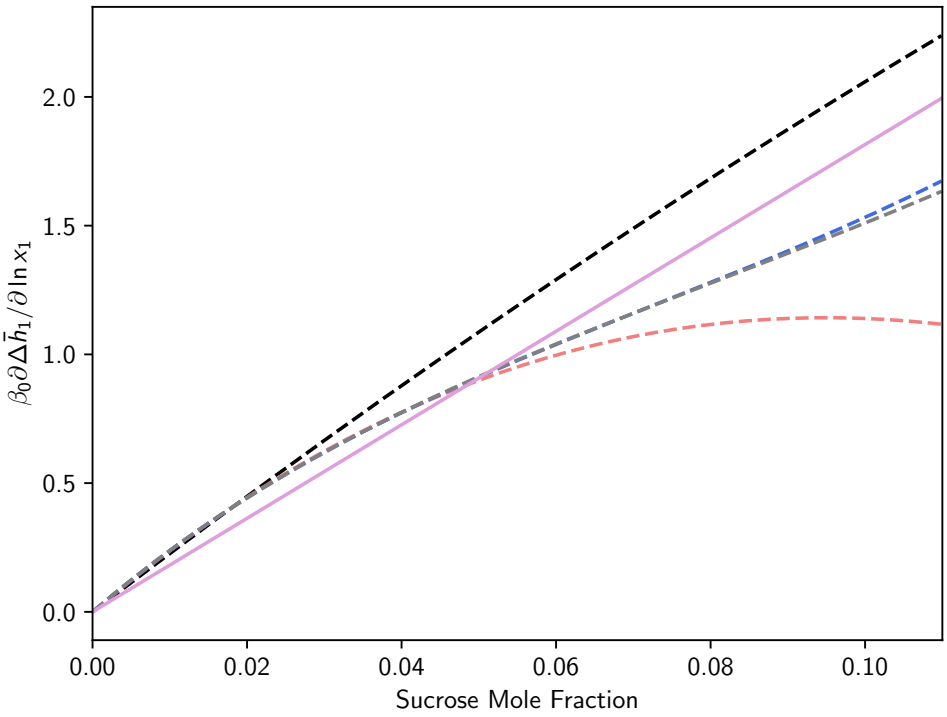


Figure 4.16: Estimates for $\beta \partial \Delta \bar{h}_1^{\mathrm{l}} / \partial \ln x_1$ derived from fits to experimental enthalpy of dilution measurements at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1, grey denoting a fit with A_1, B_1, C_1 and D_1, and purple denoting estimates from [145].

including: activity, enthalpy, heat capacity, and phase equilibria, at all available temperatures and compositions. Based on this analysis, estimates shown in Fig 4.15 and 4.16 indicate that the models derived using data at 298.15 K (and 299.15 K) with Eq. (4.15) give reasonably good agreement.

Only one report for sucrose-water solution heat capacity data at 298.15 K was found in the literature. This data is illustrated in Fig. 4.17 with the resulting best fit using Eq. (4.17). A suitable

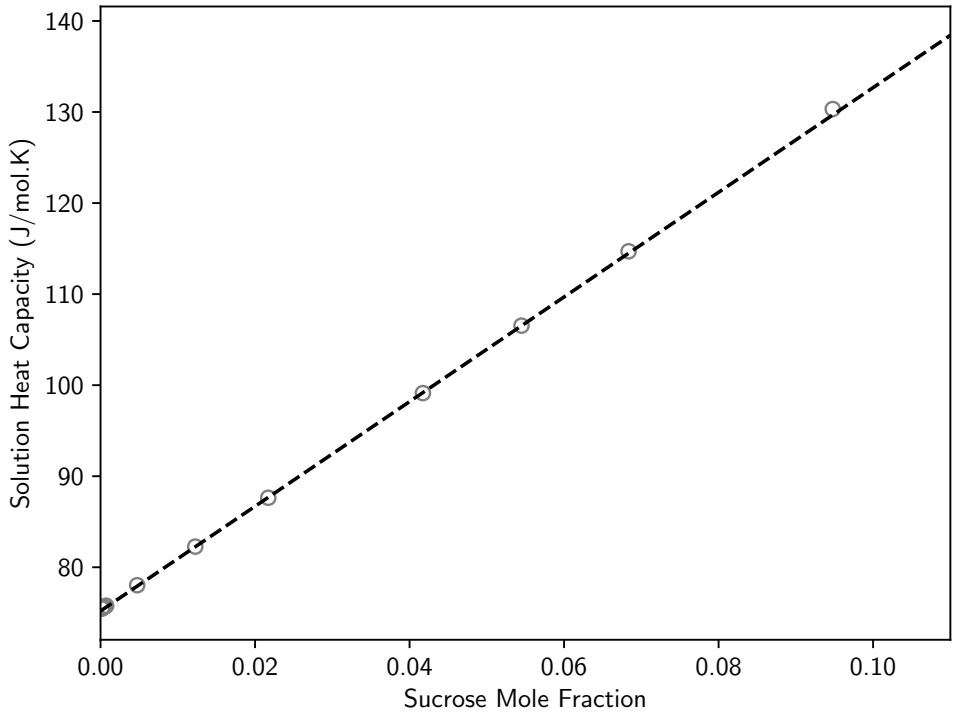


Figure 4.17: Solution heat capacity for sucrose-water solutions at 298.15 K. Open grey circles correspond to solution heat capacity measurements reported in literature [177, 178, 179, 180]. Colored dashed lines correspond to different models, with black denoting a fit with $c_{\mathrm{p},1}^{\mathrm{l},\circ}$ only.

fit to available solution heat capacity data was achieved with all parameters in the Redlich-Kister equation set to zero and varying the hypothetical pure liquid sucrose heat capacity (i.e., $c_{\mathrm{p},1}^{\mathrm{l},\circ}$), which was found to be 650 J/mol.K, giving a partial molar heat capacity at saturation (i.e., $\bar{c}_{\mathrm{p},1}^{(0)}$) with the same value. It should be noted that the molar heat capacity of pure liquid water (i.e., $c_{p,2}^{l,\circ}$) was taken as 75.2 J/mol.K. For $\Delta\bar{c}_{p,g}^{(0)}$, the pure solid solute heat capacity was taken as 420 J/mol.K for sucrose.

In a manner similar to the predictions shown for α-glycine, thermodynamic quantities derived from the above fitting process were used to make predictions for the aqueous solubility of sucrose. This included predictions generated by combining quantities derived from all models in Table 4.5

for $\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ and $\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$, all models in Table 4.6 for $\Delta \bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$ respectively, as well as the single estimate for $\Delta \bar{c}_{p,1}^{(0)}$. The results are illustrated in Fig. 4.18. Again, the predictions are split into four groups to highlight the effects of different parameter esti-

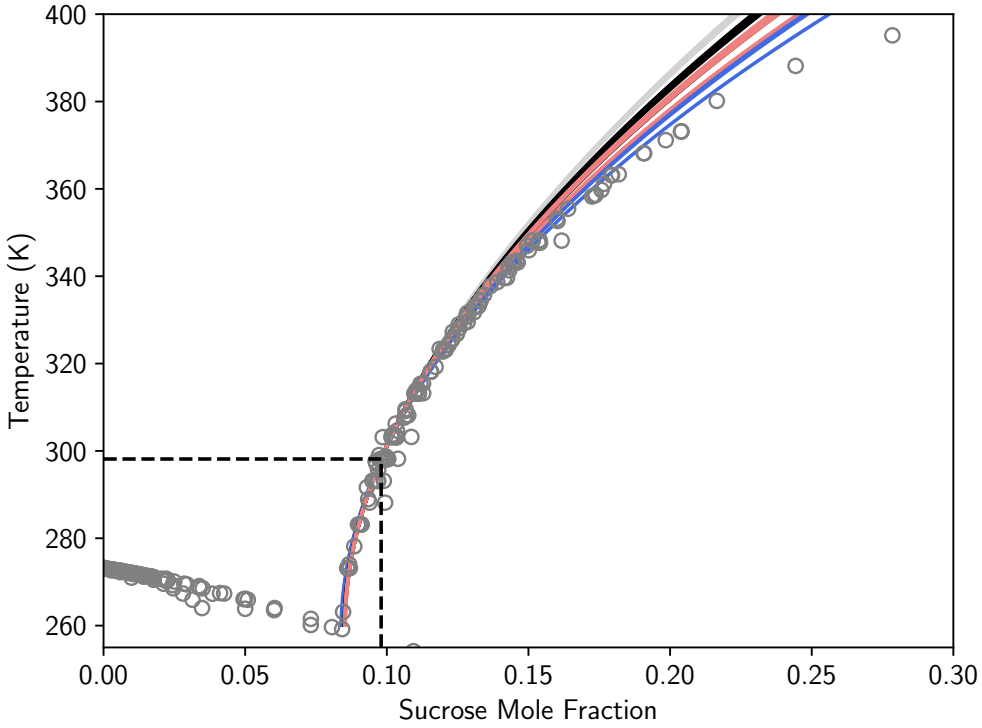


Figure 4.18: Sucrose aqueous solubility predictions based on Eq. (4.9). Open grey symbols denote directly measured solid-liquid equilibria reported in the literature (solubility [146, 147, 148, 149, 150, 151] and freezing point [181, 182, 148, 150, 183, 184]). Solid lines are predictions derived from Eq. (4.9) based on all combinations of thermodynamic quantities provided in Table 4.5 (all models), 4.6 (all models) and the single estimate of sucrose partial molar heat capacity. The blue band corresponds to all predictions based on model one in Table 4.6. The grey band corresponds to all predictions based on model two in Table 4.6. The black band corresponds to all predictions based on models three and four in Table 4.3. The red band corresponds to all predictions where $\Delta \bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$ were derived from the model described in [145].

mates. The blue group corresponds to predictions where the enthalpy parameters (i.e., $\Delta \bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$) came from model one on Table 4.6. The grey group corresponds to predictions where the enthalpy parameters came from model two in Table 4.6. The black band corrresponds to predictions where the enthalpy paramaters came from model 3 and 4 in Table 4.6. The red group corresponds to predictions where the enthalpy parameters came from the model reported in literature [145].

As shown, all predictions are in excellent agreement with each other, as well as the directly measured solubility data from 260–340 K, after which predictions in the grey and black bands appear to diverge from those in the red and blue bands, as well as direct measurements. Generally, predictions in the red band correspond to solubility estimates in excellent agreement with direct measurements up to approximately 350 K, while predictions in the blue group correspond to solubility estimates in excellent agreement with direct measurements from 260–360 K. Within each band, predictions diverge around similar temperatures driven by different estimates for $\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ and $\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$.

4.3.4 Urea

For the urea-water system, data from which the water activity coefficient could be inferred included five independent reports of urea isopiestic molality. As before, isopiestic molalities were converted to osmotic coefficeints using a model for the omsotic coefficient of reference solutes and then to solvent activity coefficient via Eq. (3.39). Experimental data (processed as $\ln \gamma_w$), with corresponding best fits derived Eq. (4.15) with parameters A_0, B_0, C_0 and D_0 are presented in Fig. 4.19. The resulting parameters for each fit are provided in Table 4.7. From Fig. 4.19, only

Table 4.7: Redlich-Kister parameter values for urea-water solvent activity data fits

Model	RSS	A_0 (-)	B_0 (-)	C_0 (-)	D_0 (-)	$\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ (-)	$\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$ (-)
1	0.07	0.69				-0.86	-0.55
2	0.0035	-1.61	1.01			0.07	0.67
3	0.0023	-0.15	-0.02	0.34		0.03	0.40
4	0.0014	0.17	0.12	-0.21	0.32	-0.06	0.03

the fit including all parameters (i.e., A_0, B_0, C_0 and D_0) produce models in good visual agreement with available data which is additionally supported by the lowest RSS in Table 4.7. Given that the only model capable of producing a fit to all available data was the most complex (i.e., all four parameters), coupled with good coverage of experimental data near saturation, an additional set of fits were generated focusing on data near the urea-water solution saturation compositon (i.e., 0.265 urea mole fraction). The motivation for this being that we're most interested in generating a good fit to data in this region, rather than the entire composition range, and trying to fit all availabe data may bias a given model away from fitting data in the region of interest required to generate accurate estimates for $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$. As before, all model forms were fit to solvent activity data; however, instead of fitting all available data, fits were resticted to

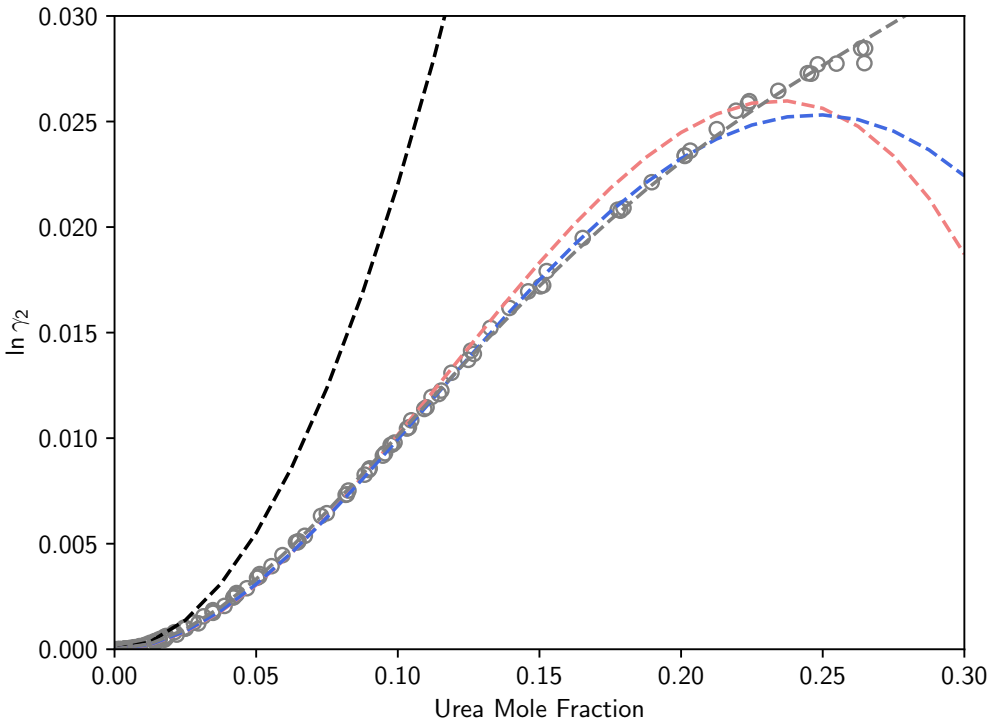


Figure 4.19: Processed water activity coefficient (Raoult's Law) data in urea-water liquid mixtures solutes at 298.15 K. Open grey symbols denote processed experimental data [185, 186, 187, 188, 189]. Colored dashed lines correspond to fits using Eq. (4.15), with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

data with a composition above 0.15 urea mole fraction. Experimental data (processed as $\ln \gamma_w$), with corresponding best fits derived Eq. (4.15) with parameters A_0, B_0, C_0 and D_0 are presented in Fig. 4.20. The resulting parameters for each fit are provided in Table 4.8. From Fig. 4.20,

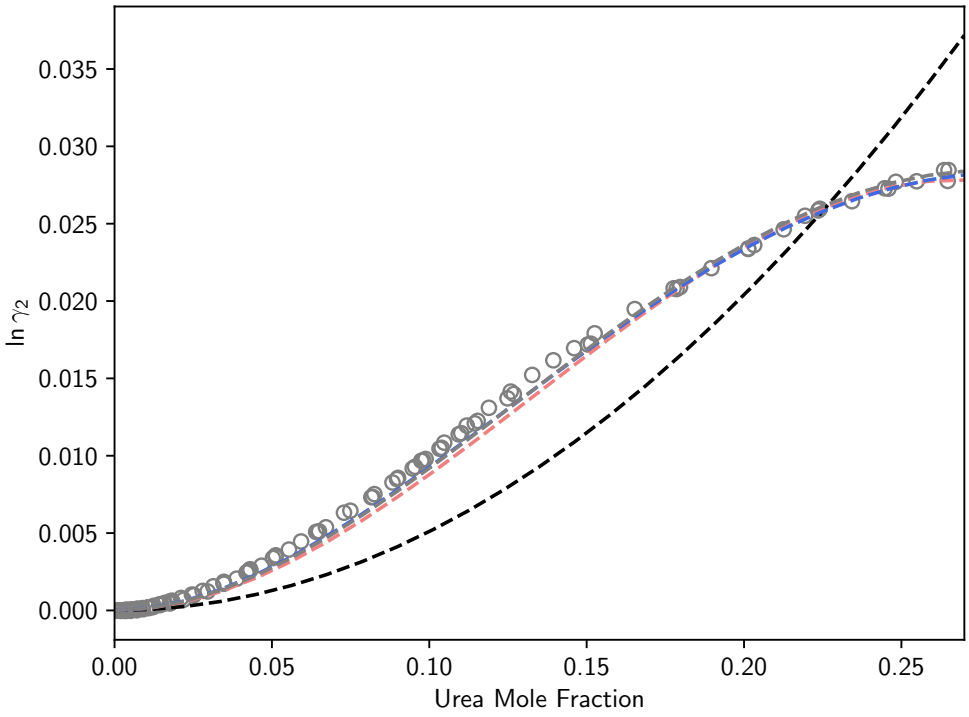


Figure 4.20: Processed water activity coefficient (Raoult's Law) data in urea-water liquid mixtures solutes at 298.15 K.

Open grey symbols denote processed experimental data [185, 186, 187, 188, 189]. Colored dashed lines correspond to fits using Eq. (4.15), with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

Table 4.8: Redlich-Kister parameter values for urea-water solvent activity data fits

Model	RSS	A_0 (-)	B_0 (-)	C_0 (-)	D_0 (-)	$\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ (-)	$\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$ (-)
1	4.6e-3	0.51				-0.20	-0.13
2	2.3e-5	-1.02	0.73			-0.004	0.45
3	1.18e-5	0.024	-0.047	0.291		-0.02	-0.29
4	1.4e-5	0.044	-0.106	0.39	-0.056	-0.016	0.36

restricting the data used during fitting made it possible to easily find parameter sets for models with two and three parameters in excellent agreement with available data near the saturation composition (i.e., 0.265 urea mole fraction), when compared to the fits to all data illustrated in Fig. 4.19 and summarised in Table 4.7. However, the fit with only one parameter was still unable to describe any of the data, indicating the model structure is not suitable for this system.

Comparing Fig. 4.19 and 4.20, there are four models that give a reasonable description of the solvent activity data for urea-water liquid mixtures at 298.15 K; model 4 in Table 4.7 and models 2–4 in Table 4.8. The first and second solute activity coefficient partial derivatives with respect to composition derived from each of these fits to solvent activity data are presented in Fig. 4.21 and 4.22, respectively. As shown in Fig. 4.21 and 4.22, the partial derivatives (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$

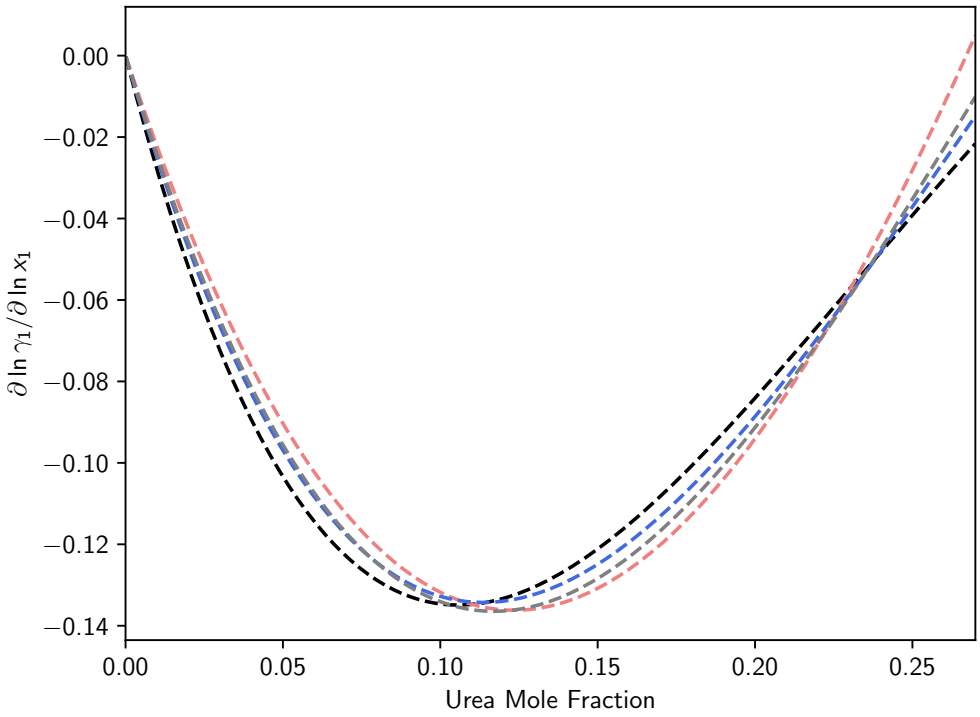


Figure 4.21: Solute activity coefficient first partial derivative with respect to solute composition estimated from best fits shown in Fig. 4.19. Colored dashed lines correspond to different models, with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) are in fairly good agreement; however, all of which begin to diverge close to the urea saturation composition.

For urea, as with α-glycine and sucrose, there exist reports of the enthalpy of dilution at 298.15 K which can, in theory, be used to estimate $\bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$, as long as at least

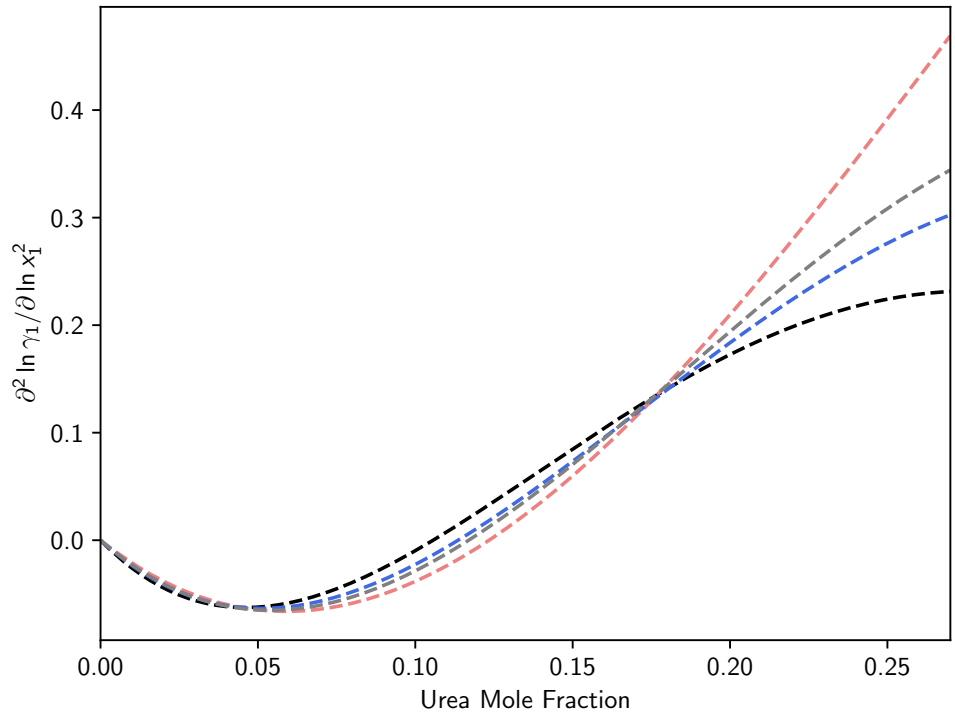


Figure 4.22: Solute activity coefficient first partial derivative with respect to solute composition estimated from best fits shown in Fig. 4.19. Colored dashed lines correspond to different models, with black denoting a fit with A_0, red denoting a fit with A_0 and B_0, blue denoting a fit with A_0, B_0 and C_0 and grey denoting a fit with A_0, B_0, C_0 and D_0.

one enthalpy of solution measurement is available. However, a study reporting values of the enthalpy of solution from infinite dilution to saturation was also found. It should be noted that no other data sets reporting enthalpy of solution measurements at solution concentrations greater than 'infinitely dilute' were found to validate this data set. The enthalpy of solution was interpreted via:

$$\begin{aligned}\Delta H^{\text{sol.}} &= n_1\left(h_1^{\text{l},\circ} + \bar{h}_1^{\text{E}}\right) + n_2\left(h_2^{\text{l},\circ} + \bar{h}_2^{\text{E}}\right) - n_1 h_1^{\text{s},\circ} - n_2 h_2^{\text{l},\circ} \\ &= n_1\left(\Delta h_1^{\infty,s} - h_1^{\text{l},\infty}\right) + n_1\bar{h}_1^{\text{E}} + n_2\bar{h}_2^{\text{E}}\end{aligned} \tag{4.27}$$

where $\Delta h_1^{\infty,s}$ denotes the enthalpy of solution to infinite dilution (taken as 15,400 J/mol) and $h_1^{\text{l},\infty}$ denotes the molar enthalpy of solute in solution at infinite dilution (i.e., infinitely dilute reference state c.f. Henry's Law). All other quantities have been identified previously and retain those definitions.

Initially, fitting was completed by aggregating both types of data (i.e., solution and dilution); however, given challenges already discussed with visualising fits to dilution data, a set of models were generated by fitting each type of data independently. The models generated by fitting only enthalpy of dilution data are shown in Fig 4.23 while models generated by fitting enthalpy of solution data are shown in Fig. 4.24, both of which predicting the enthalpy of solution data, as it is possible to readily visualise this.

Parameters derived from fitting enthalpy of dilution data are provided in Table 4.9, while parameters derived from fitting enthalpy of solution data are provided in Table 4.10. The

Table 4.9: Redlich-Kister parameter values for urea enthalpy of dilution data fits

Model	RSS	A_1 (J/mol)	B_1 (J/mol)	C_1 (J/mol)	D_1 (J/mol)	$\Delta\bar{h}_1^{(0)}$ (J/mol)	$\beta_0\partial\Delta\bar{h}_1^{(0)}/\partial\ln x_1$ (-)
1	1982213	4.80				9975.9	-1.86
2	438570	-10.31	5.73			12660.62	0.97
3	398015	-0.21	0.52	0.99		12393.1	0.49
4	145107	0.025	-0.064	0.54	0.68	9339.7	-4.42

enthalpy of solution derived from each fit is shown in Fig. 4.23 and 4.24 alongside experimental data, for fits derived from dilution and solution measurements, respectively.

Estimates for $\bar{h}_1^l - h_1^{\text{l},\infty}$ are presented in Fig. 4.26 and estimates for $\beta_0\partial\Delta\bar{h}_1/\partial\ln x_1$ are presented in Fig. 4.28. As shown, generally speaking, there is disagreement between each model for all enthalpy properties, leading to potentially significant difference in $\Delta\bar{h}_1^{(0)}$ and $\partial\Delta\bar{h}_1^{(0)}/\partial\ln x_1$. It is interesting to note that model 2 in Table 4.10 gives the best approximation of the enthalpy of solution data, compared to all other models.

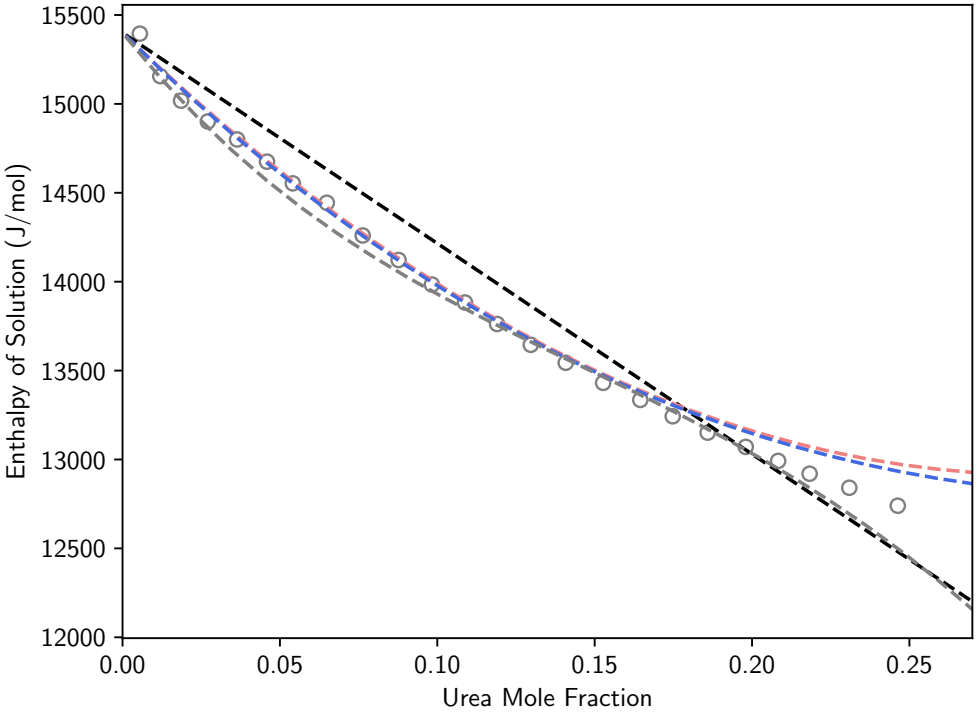


Figure 4.23: Estimates for urea-water enthalpy of solution derived from fits to experimental enthalpy of dilution measurements [190, 191] for urea-water mixtures at 298.15 K using Eq. (4.15). Open grey symbols correspond to experimental measurements [192]. Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

Table 4.10: Redlich-Kister parameter values for urea enthalpy of solution data fits

Model	RSS	A_1 (J/mol)	B_1 (J/mol)	C_1 (J/mol)	D_1 (J/mol)	$\Delta\bar{h}_1^{(0)}$ (J/mol)	$\beta_0\partial\Delta\bar{h}_1^{(0)}/\partial \ln x_1$ (-)
1	692605	4.9				9839.8	-1.91
2	24426	-8.1	4.9			12124.3	0.46
3	21115	-0.89	0.23	1.42		11812.6	0.04
4	19806	0.12	0.09	0.82	0.36	11718.8	-0.15

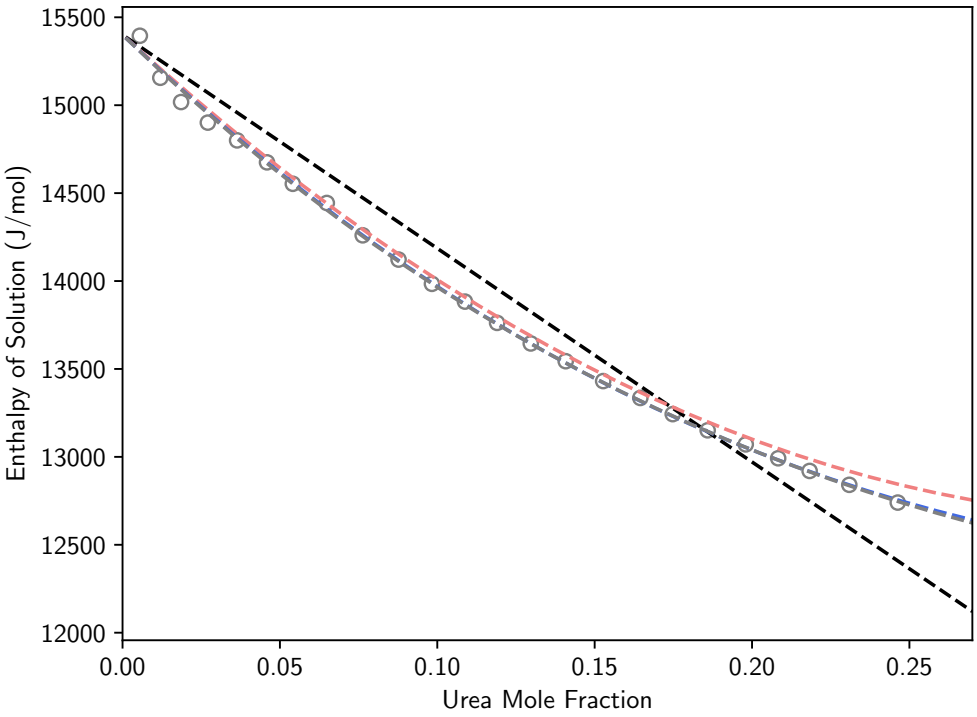


Figure 4.24: Estimates for urea-water enthalpy of solution derived from fits to experimental enthalpy of solution measurements [192] for urea-water mixtures at 298.15 K using Eq. (4.15). Open grey symbols correspond to experimental measurements [192]. Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

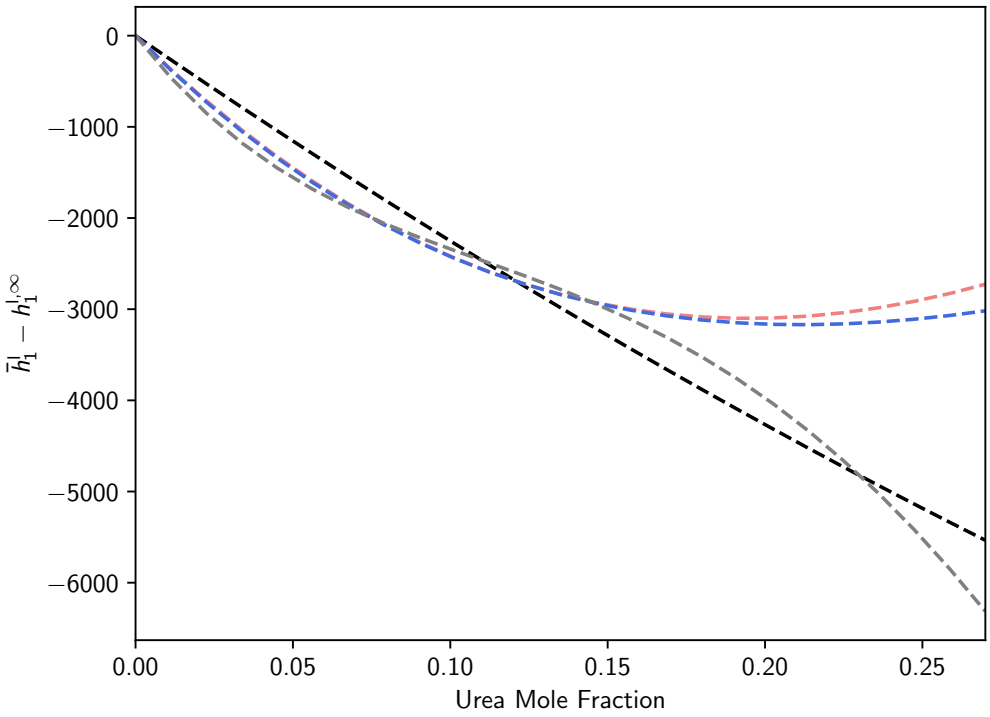


Figure 4.25: Estimates for $\bar{h}_1^{\mathrm{l}} - h_1^{\mathrm{l},\infty}$ derived from fits to experimental enthalpy of dilution measurements for urea-water mixtures at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

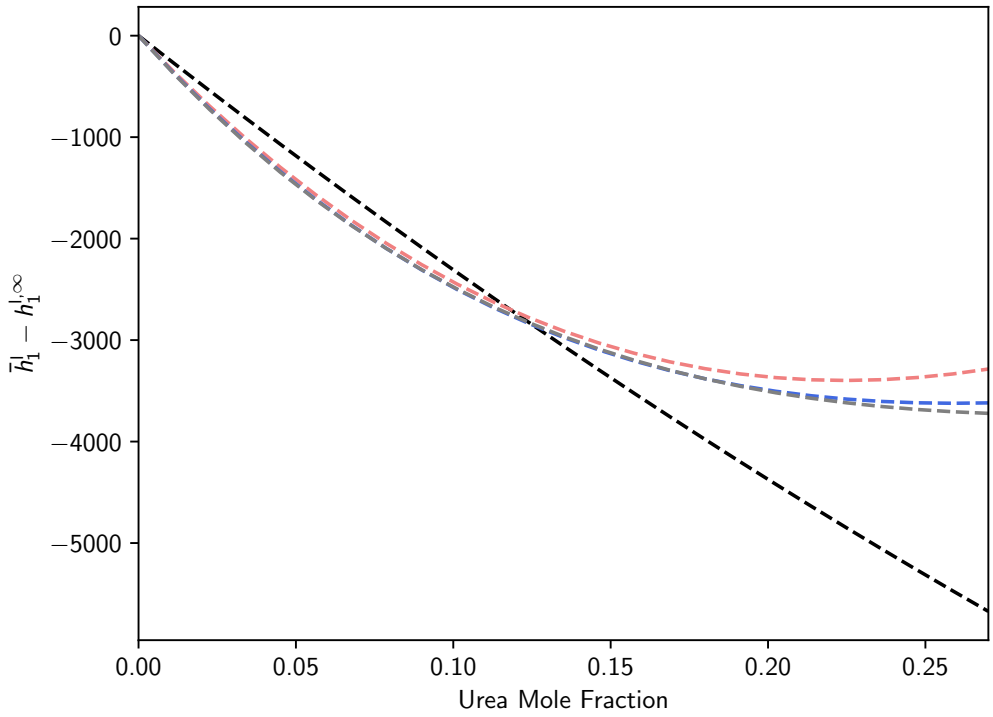


Figure 4.26: Estimates for $\bar{h}_1^{\mathrm{l}} - h_1^{\mathrm{l},\infty}$ derived from fits to experimental enthalpy of dilution measurements for urea-water mixtures at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

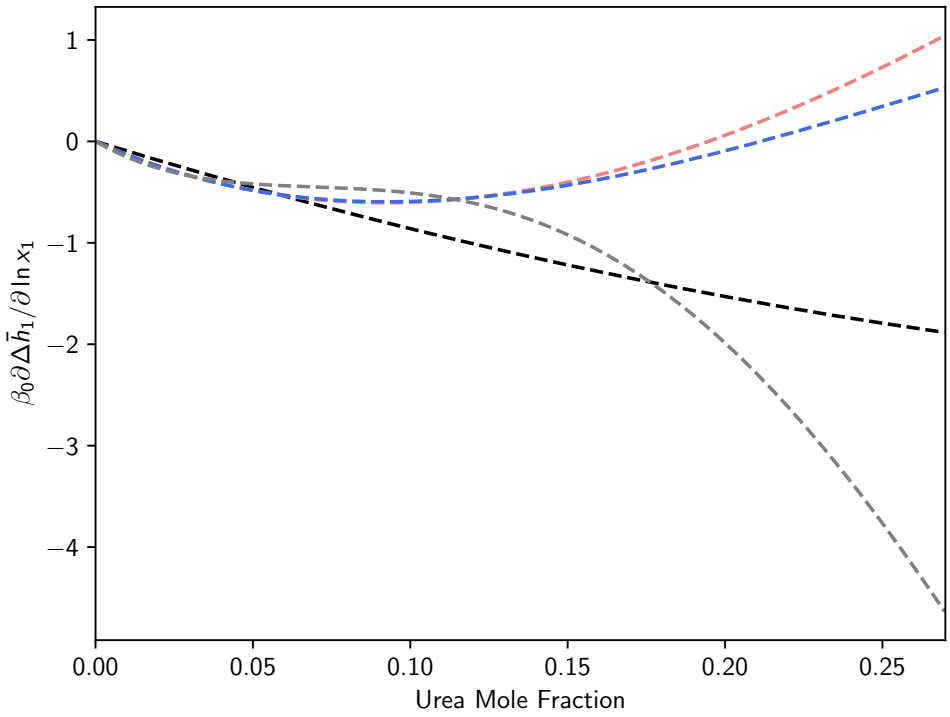


Figure 4.27: Estimates for $\beta \partial \Delta \bar{h}_1^{\mathrm{l}} / \partial \ln x_1$ derived from fits to experimental enthalpy of dilution measurements at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

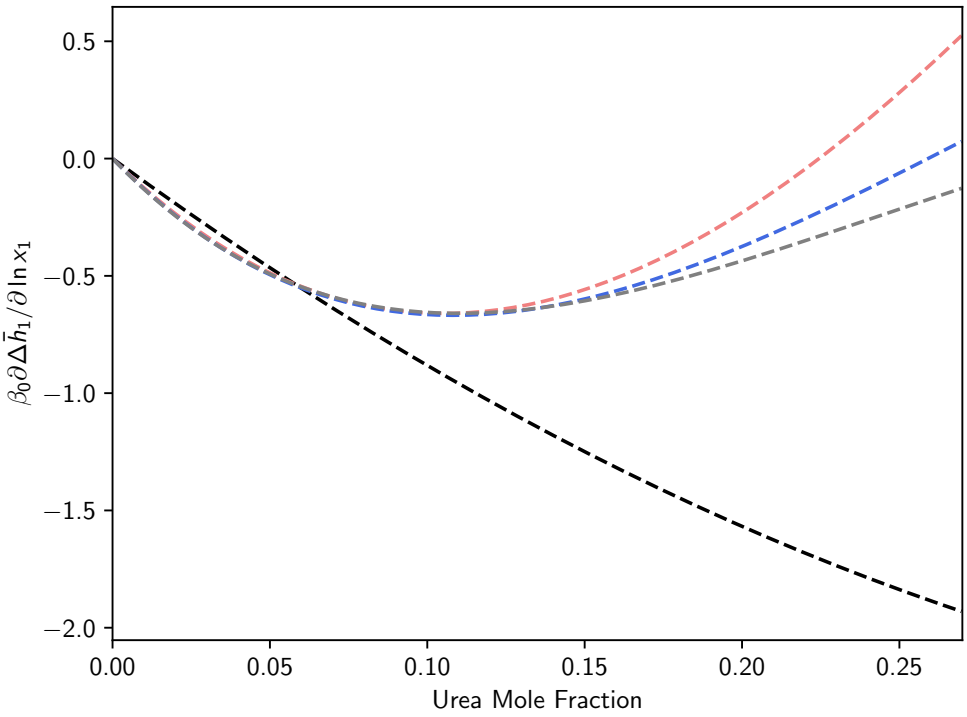


Figure 4.28: Estimates for $\beta \partial \Delta \bar{h}_1^{\mathrm{l}} / \partial \ln x_1$ derived from fits to experimental enthalpy of dilution measurements at 298.15 K using Eq. (4.15). Colored dashed lines correspond to different models, with black denoting a fit with A_1, red denoting a fit with A_1 and B_1, blue denoting a fit with A_1, B_1 and C_1 and grey denoting a fit with A_1, B_1, C_1 and D_1.

Heat capacity data for urea-water solutions reported in the literature at 298.15 K is shown in Fig. 4.29 with best fits derived from Eq. (4.17) where the molar heat capacity of pure liquid water (i.e., $c_{p,2}^{l,\circ}$) was taken as 75.2 J/mol.K. The resulting parameter values for each fit are provided in Table 4.11. Models and experimental data are presented in Fig. 4.29 and the urea

Table 4.11: Redlich-Kister parameter values and for urea-water solution heat capacity fits

Model	RSS	A_2 (J/mol.K)	B_2 (J/mol.K)	C_2 (J/mol.K)	D_2 (J/mol.K)	$c_{\mathrm{p},1}^{\mathrm{s},\circ}$ (J/mol.K)	$\Delta\bar{c}_{\mathrm{p},1}^{(0)}$ (J/mol.K)
1	1.24	-95.38				191.25	-51.58
2	0.56	267.61	-155.23			-24.09	-65.79
3	0.56	465.05	-290.86	44.22		-129.64	-66.9
4	0.56	243.23	-116.94	-42.21	18.57	-14	-66.27

partial molar heat capacity is presented in Fig. 4.30. As shown, all models provide plausible fits to available solution heat capacity data. However, as noted for several properties already, this apparent similarity is lost when thermodynamic quantities required to make solubility predictions are derived, which is the case for estimates of the partial molar heat capacity of urea. Estimates for $\Delta\bar{c}_{p,1}^{(0)}$ from each model are provided in Table 4.11, where the heat capacity of pure crystalline urea at 298.15 K was taken as 192.3 J/mol.K.

As with α-glycine and sucrose, thermodynamic quantities derived from the above fitting process were used to make predictions for the aqueous solubility of urea. This included predictions generated by combining quantities derived from model 4 in Table 4.7 and model 2–4 in Table 4.8 for $\partial \ln \gamma_1^{(0)} / \partial \ln x_1$ and $\partial^2 \ln \gamma_1^{(0)} / \partial \ln x_1^2$, as well as all models in Table 4.11 for $\Delta\bar{c}_{\mathrm{p},1}^{(0)}$. For parameters related to the enthalpy, solubility predictions have been split based on enthalpy parameters (i.e., $\Delta\bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta\bar{h}_1^{(0)} / \partial \ln x_1$) derived from the enthalpy of dilution data and the enthalpy of solution fits and are illustrated in Fig. 4.31 and 4.32, respectively.

As shown, for both sets of predictions (i.e., those derived with models parameterised using enthalpy of dilution data, as well as those with enthalpy of solution data), there are subsets of predictions in excellent agreement with directly measured solubility data. Focusing specifically on predictions made with models derived from enthalpy of dilution data (i.e., Fig. 4.31), the black and blue bands corresponds to predictions made with model 1 and 4 in Table 4.9, respectively, and yield predictions in poor agreement with experimental data at all temperatures and compositions. These predictions fail to predict the local slope around the chosen reference point (i.e., 298.15 K and 0.265 urea mole fraction), indicating these models are incorrect. The red and grey bands correspond to model 2 and 3 in Table 4.9, and many of the model combinations yield predictions in excellent agreement with directly measured solubility data, with the exception of two predictions

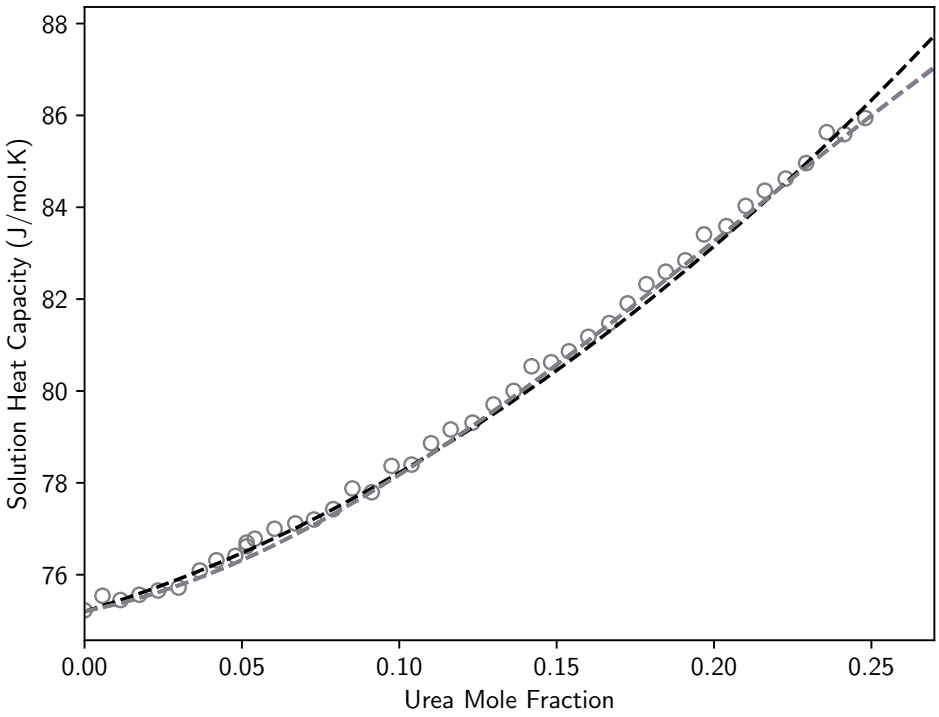


Figure 4.29: Solution heat capacity for urea-water solutions at 298.15 K. Open grey circles correspond to solution heat capacity measurements reported in literature [193, 194, 195, 196]. Colored dashed lines correspond to different models, with black denoting a fit with A_2, red denoting a fit with A_2 and B_2, blue denoting a fit with A_2, B_2 and C_2 and grey denoting a fit with A_2, B_2, C_2 and D_2.

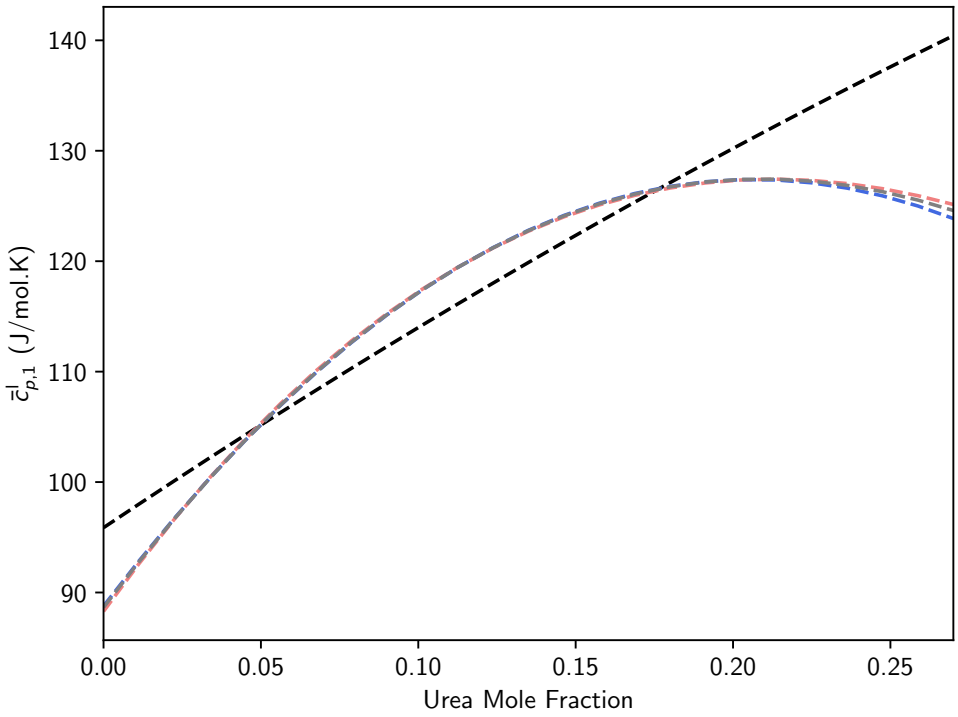


Figure 4.30: Partial molar heat capacity of urea in urea-water mixtures at 298.15 K derived from fits to soution heat capacity data using Redlich-Kister models. Colored dashed lines correspond to different models, with black denoting a fit with A_2, red denoting a fit with A_2 and B_2, blue denoting a fit with A_2, B_2 and C_2 and grey denoting A_2, B_2, C_2 and D_2.

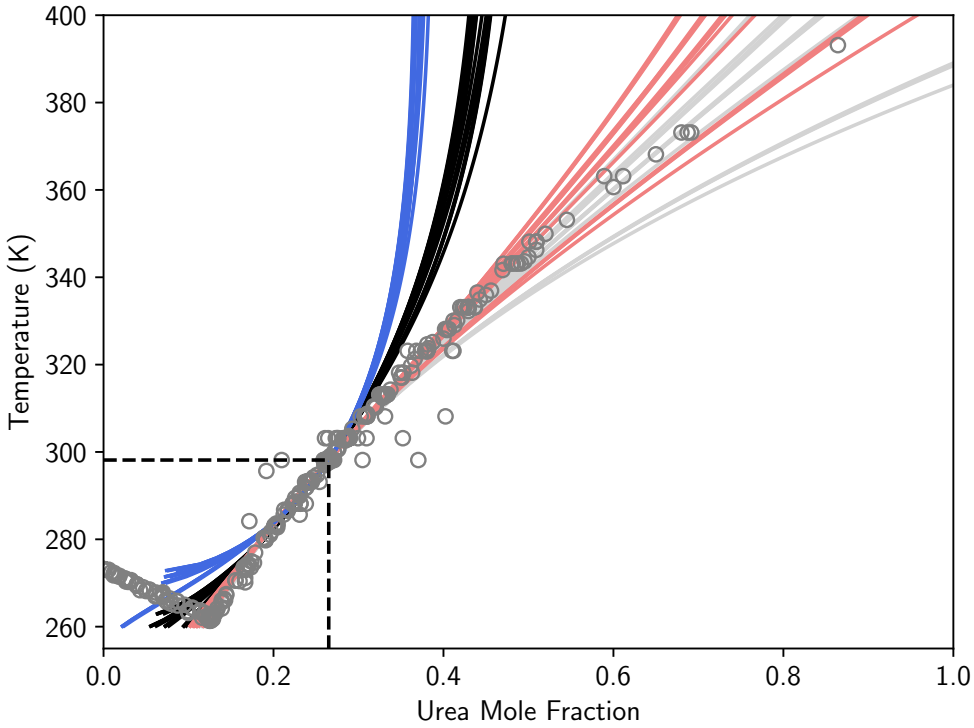


Figure 4.31: Urea aqueous solubility predictions based on Eq. (4.9). Open grey symbols denote directly measured solid-liquid equilibria reported in the literature (solubility [152, 153, 154, 155, 156, 157, 158] and freezing point [197]). Solid lines are predictions derived from Eq. (4.9) based on all combinations of thermodynamic quantities provided in Table 4.7 (model four), 4.8 (models two–four), 4.9 (all models) and 4.11 (all models). The black band corresponds to all predictions based on model one in Table 4.9. The red band corresponds to all predictions based on model two in Table 4.9. The grey band corresponds to all predictions based on models three in Table 4.9. The blue band corresponds to all predictions based on models four in Table 4.9.

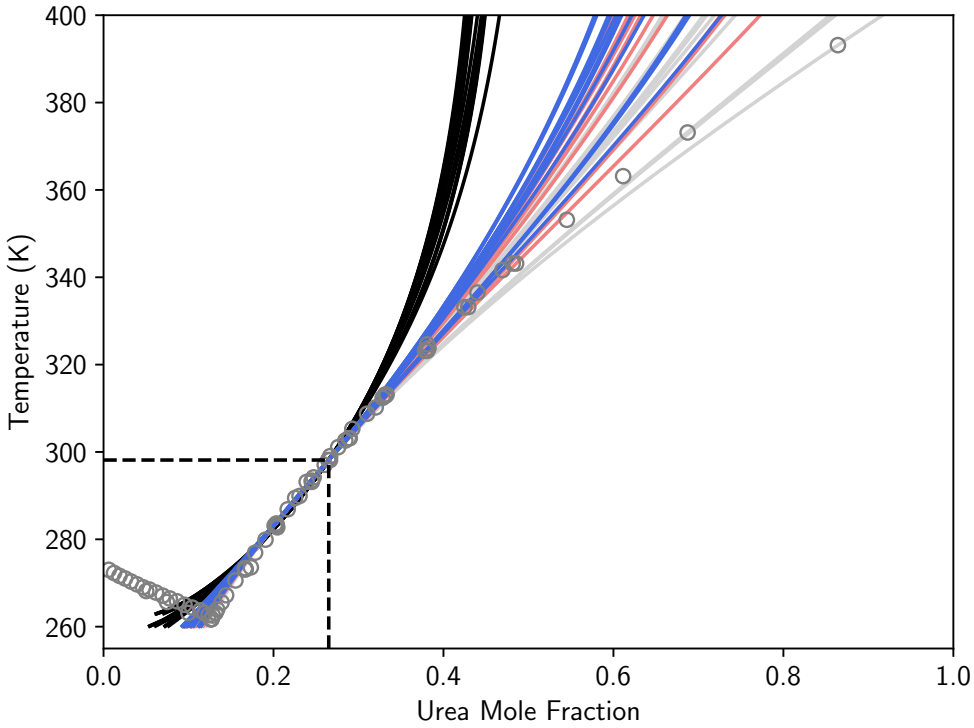


Figure 4.32: Urea aqueous solubility predictions based on Eq. (4.9). Open grey symbols denote directly measured solid-liquid equilibria reported in the literature (solubility [152, 153, 154, 155, 156, 157, 158] and freezing point [197] Solid lines are predictions derived from Eq. (4.9) based on all combinations of thermodynamic quantities provided in Table 4.7 (model four), 4.8 (models two–four), 4.10 (all models) and 4.11 (all models). The black band corresponds to all predictions based on model one in Table 4.10. The grey band corresponds to all predictions based on model two in Table 4.10. The red band corresponds to all predictions based on models three in Table 4.10. The blue band corresponds to all predictions based on models four in Table 4.10.

in the grey band which diverge from experimental data far quicker than all other models—it should be noted that these are associated to model 4 in Table 4.7 used to estimate activity parameters (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$). A simplified view of predictions is provided in Fig. 4.33, which shows the range of predictions based on the most plausible models.

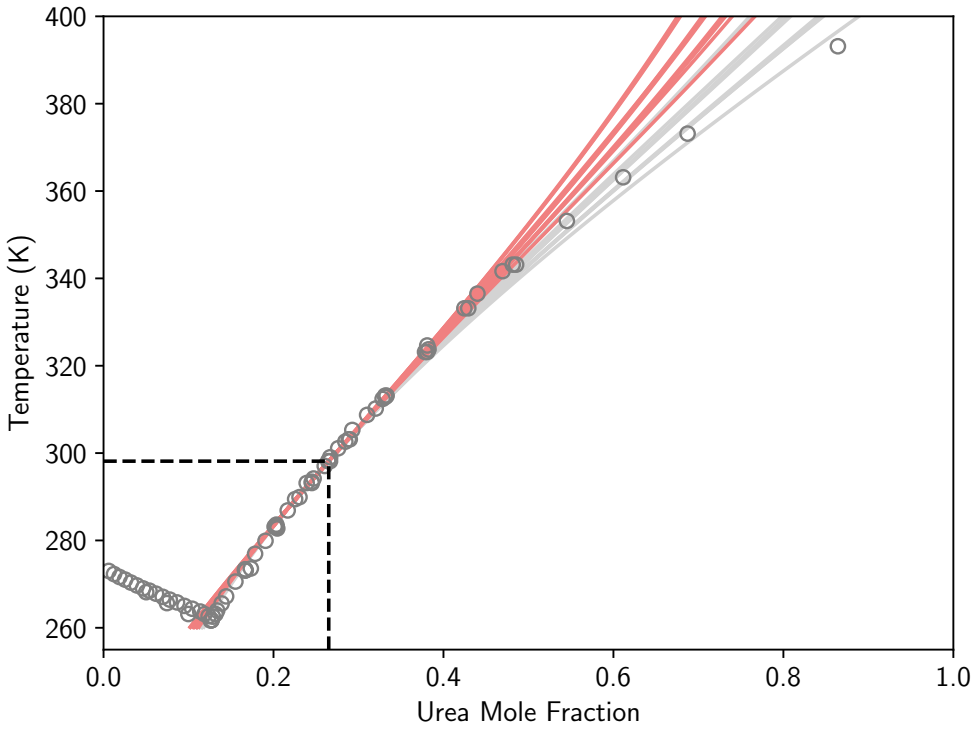


Figure 4.33: Urea aqueous solubility predictions based on Eq. (4.9). Open grey symbols denote directly measured solid-liquid equilibria reported in the literature (solubility [152, 153, 154, 155, 156, 157, 158] and freezing point [197] Solid lines are predictions derived from Eq. (4.9) based on all combinations of thermodynamic quantities provided in Table 4.8 (models two–four), 4.9 (models two and three) and 4.11 (all models). The red band corresponds to all predictions based on model two in Table 4.9. The grey band corresponds to all predictions based on models three in Table 4.9.

Focusing on predictions made with models derived from enthalpy of solution data (i.e., Fig. 4.32), the black band corresponds to model 1 in Table 4.10 and, in a manner similar that observed with the one parameter model for dilution data, predicts solubility in poor agreement with available data. Comparing to Fig. 4.24, this model failed to describe available enthalpy of solution data, potentially indicating these predictions should be ignored. The grey, red and blue bands correspond to predictions derived from models 2–4 in Table 4.10 and lead to predictions in better agreement with solubility measurements. Generally speaking, all parameter combinations lead to predictions

in excellent agreement data from 260–340 K, before diverging from each other, and experimental data. In a manner similar to that shown for predictions based on models parameterised using enthalpy of dilution data only, removing predictions where the activity parameters (i.e., $\partial \ln \gamma_1 / \partial \ln x_1$ and $\partial^2 \ln \gamma_1 / \partial \ln x_1^2$) are taken from model 4 in Table 4.7 are shown in Fig. 4.34. By comparison

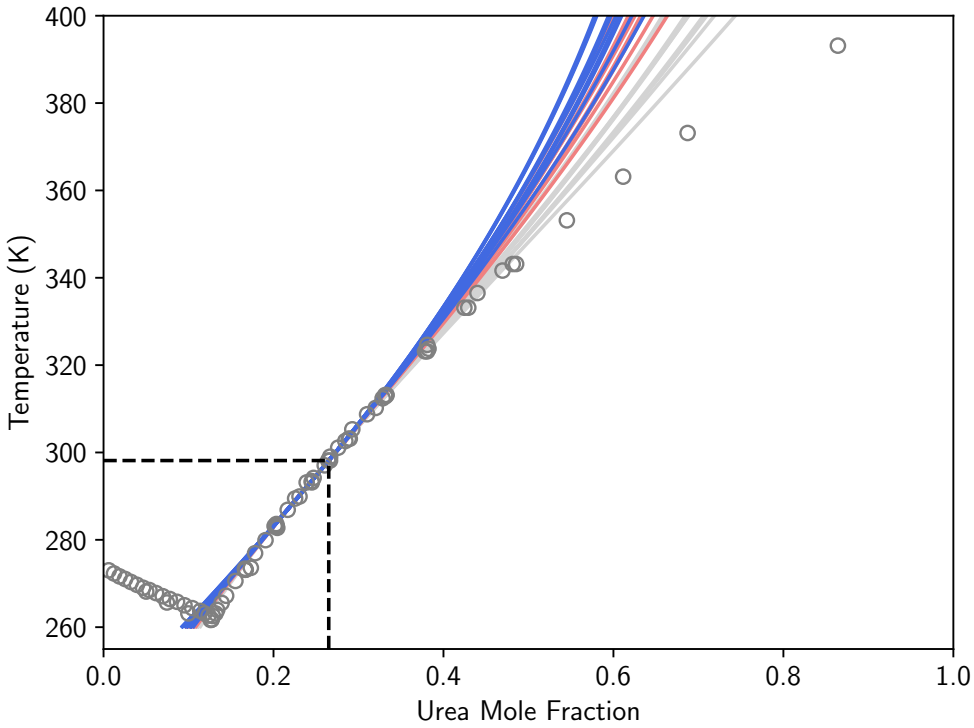


Figure 4.34: Urea aqueous solubility predictions based on Eq. (4.9). Open grey symbols denote directly measured solid-liquid equilibria reported in the literature (solubility [152, 153, 154, 155, 156, 157, 158] and freezing point [197] Solid lines are predictions derived from Eq. (4.9) based on all combinations of thermodynamic quantities provided in Table 4.8 (models two–four), 4.10 (models two–four) and 4.11 (all models). The grey band corresponds to all predictions based on model two in Table 4.10. The red band corresponds to all predictions based on models three in Table 4.10. The blue band corresponds to all predictions based on models four in Table 4.10.

to Fig. 4.32, predictions based on these parameters are in better agreement with experimental data across the full range of temperatures.

4.4 Conclusions

In this chapter, an approach allowing extrapolation of temperature dependent solubility by considering the chemical potential of the solid in both solid and liquid solution in terms of Taylor series in

terms of $\ln x_g$ and β, truncated at the 2nd-order was presented. Like conventional thermodynamic interpretations of solubility, our proposed model retains a key advantage for predictive purposes in that solubility is expressed in terms of thermodynamic quantities related to experimentally measurable properties. Crucially, however, none of the data required to make solubility predictions include the pure solute melting properties, which can be challenging to measure and limits the applicability of conventional approaches.

The aqueous solubility of α-glycine, sucrose and urea was predicted using thermodynamic quantities derived from available experimental data reported at 298.15 K. For all systems, solubility predictions based on the new approach give estimates that are sensitive to the data and modeling approach used to derive required quantities. This was particularly evident for α-glycine and urea, where some predictions diverged rapidly from available direct measurement data. As shown, thermodynamic quantities required to make extrapolations were estimated by correlating measurable properties with a mathematical model, then applying mathematical manipulations (e.g., partial differentiation, etc), which brings several sources of uncertainty usually associated to any modeling application. Small differences in functional form may lead to meaningful differences in extrapolated solubility far from the chosen reference point, which is particularly problematic for systems where data required to estimate relevant thermodynamic quantities are limited, making model selection and parameterisation challenging.

However, it was found that certain interpretations of experimental data gave predictions in good agreement with direct solubility measurements. This was most convincing for the sucrose-water system, where multiple modelling approaches gave similar results, likely giving the most robust assessment of the modelling approach. Across all systems, predictions were in excellent agreement with direct measurements between 260–340 K regardless of the modelling approach taken, when plausible fits were used. Above this, predictions started to diverge amongst themselves, as well as from the direct solubility measurements. Interestingly, this agrees with findings noted in Chapter 3, and may indicate an upper limit in the reliability of Eq. (4.9), when relevant thermodynamic quantities are derived via mathematical modeling of experimental data at 298.15 K.

Generally speaking, the most uncertain parameters for all systems were $\Delta \bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$, driven by limited availability of enthalpy data from which the models that yield these quantities can be parameterised. With this in mind, it would be advantageous to perform additional measurements with a particular focus on quantities from which $\Delta \bar{h}_1^{(0)}$ and $\beta_0 \partial \Delta \bar{h}_1^{(0)} / \partial \ln x_1$ can be derived with greater certainty—this may involve measuring the enthalpy of dilution, or solution as shown for Urea.

Chapter 5

Beyond Approximate Approaches

5.1 Introduction

In Chapter 3 and 4, approximate methods allowing estimation of solubility as a function of temperature were presented and benchmarked on organic binary solute-solvent systems. Both methods made it possible to 'predict' solubility by evaluating thermodynamic quantities other than solubility at a single point on the solubility curve, alongside a single estimate for solubility. However, many of these quantities are not widely reported in the literature and were estimated by considering alternative measureable quantities that can be related to the solute-solvent solution and pure crystalline solute phases. These approaches might be advantageous in situations where availability of data is limited to a single temperature (i.e., 298.15 K) or the level of predictive accuracy is less important (particularly at temperatures far from the chosen reference temperature at which relevant thermodynamic quantities are evaluated); however, they are approximate through introduction of truncated Taylor series expansions.

In this chapter, the general problem of constructing Gibbs free energy models as functions of temperature and composition by mathematically modelling experimental data is reviewed. As described, this makes it possible to estimate solubility to a level limited only by our ability to collect and model reliable experimental data—no approximations are required, in principle. This general approach is then examined by estimating the temperature dependent solubility of sucrose in water by developing explicit representations for the Gibbs free energy of pure crystalline sucrose as a function of temperature, and the Gibbs free energy of the sucrose-water liquid mixture as a function of temperature and composition, using a broad range of experimental data reported in the literature. For a specific interpretation of the Gibbs free energy of pure components and binary mixtures, it is shown that estimating solubility from the eutectic point to the pure crystal-vapor-liquid triple point is possible, with only a single estimate of solubility, retaining one of the key features of the approaches described in Chapter 3 and 4, but without the need to introduce

any approximations or limiting the experiments used during modelling.

The rest of this chapter is structured as follows. First, a discussion of modeling thermodynamic properties and phase equilibria is provided, expanding that presented in Chapter 2. Then, some illustrative applications of this theory are discussed, including the introduction and summary of the so-called CALPHAD method—a systematic approach to the construction of phase diagrams widely applied in the literature [198]. With the CALPHAD method introduced, the relationship between functional form, parameter values and a set of common experimental measurements is reviewed. A key finding from this analysis is the illustration that estimating solubility using conventional CALPHAD models requires a minimum set of experimental data such that all model parameters can be evaluated. Crucially, it is shown that only a single measure of solubility is required so that all model parameters can be found, retaining one of the key features of the approaches presented in Chapters 3 and 4, and the primary objective of this thesis (c.f. Chapter 1).

After this, a review of experimental data reported for the sucrose-water system is presented, supplementing that provided in Chapter 4 which focused on experimental data reported at 298.15 K, which is followed by a discussion of different thermodynamic modeling studies performed on the sucrose-water system. Then, the sucrose-water data is modelled following the approach applied during a CALPHAD assessment, where single component phases are modelled explicitly as temperature expansions and the binary liquid mixture is modelled with a Redlich-Kister expansion for the excess contribution (defined relative to Raoult's Law). Interestingly, we find that it is possible to 'predict' the solubility curve in excellent agreement with available experimental data with a simple model for the binary liquid mixture, despite the underlying model being in significant disagreement with differential enthalpy of dilution data that has been used to develop models previously published in the literature.

5.2 Simultaneous Representation of Thermodynamic Data

As described in Chapter 2, one relationship characterising the equilibrium between a solid and a liquid (free to exchange energy and matter) is:

$$\beta\mu_1^{\mathrm{s},\circ}(T) = \beta\mu_1^{\mathrm{l,sat}}(T, x_1(T)) \tag{5.1}$$

where $\mu_1^{\mathrm{s},\circ}$ and $\mu_1^{\mathrm{l,sat}}$ are the chemical potential of solute in crystalline and liquid solution phases, respectively maintaining the same notation as used in previous chapters. If a parameterised functional form for both chemical potentials was known, it would be possible to solve this relationship and find the equilibrium solution composition at chosen temperature, pressure and fixed total amount of each component. Assuming the chemical potential functions were reliable at all temperatures and solution compositions, this gives an approach to estimate solubility. However, as

noted already, it is not possible to measure the chemical potential directly and so alternative approaches leveraging this relationship are required, motivating the work presented in Chapter 3 and 4.

Although potentially useful in certain situations, the approaches presented in Chapter 3 and 4 are necessarily approximate given that predictions are made possible by introduction of truncated Taylor series expansions. Furthermore, although not strictly necessary, evaluation of thermodynamic quantities required to make predictions is conveniently approached by considering experimental measurements completed at a chosen reference temperature. This is potentially problematic because practitioners looking to estimate solubility might have sufficient resources to make a diverse set of measurements across a broad range of temperature and composition in excess of the single temperature approach described in Chapter 3 and 4.

For simple systems comprising a single-component solute and solvent, such as those studied in Chapter 3 and 4, the solubility curve is just one phase boundary in the whole phase diagram, the others being a boiling and freezing curve, as illustrated in Fig. 5.1. It should be noted that the complete phase diagram for a binary system will extend to temperatures above and below the range illustrated; however, for the purpose of this chapter these regions are not considered. For the phase diagram illustrated in Fig. 5.1, the stable phases include: pure crystalline component 1, pure crystalline component 2, pure liquid component 2, pure vapor component 2, as well as an components 1 and 2 in a liquid mixture. Each phase boundary is related to the solute-solvent solution, via:

$$\beta\mu_2^{\mathrm{v},\circ} = \beta\mu_2^{\mathrm{l}} \tag{5.2}$$

$$\beta\mu_1^{\mathrm{s},\circ} = \beta\mu_1^{\mathrm{l}} \tag{5.3}$$

$$\beta\mu_1^{\mathrm{s},\circ} = \beta\mu_1^{\mathrm{l}}, \tag{5.4}$$

for the boiling, solubility and freezing point curves, respectively, where 1 denotes solute and 2 denotes solvent and the superscript $\circ$ denotes a pure phase. Like solubility, it is possible to develop explicit equations for both the freezing and boiling point curves. For example, the temperature at which a solute-solvent liquid mixture will freeze as a function of temperature at fixed pressure (i.e., solvent freezing curve) can be expressed as:

$$\begin{aligned}\ln\left(1 - x_1(T)\right) = -\frac{\Delta h_2^{\mathrm{fus.}}(T_m)}{RT_m}\left(\frac{T_m}{T} - 1\right) + \int_{\beta(T_m)}^{\beta(T)}\int_{T_m}^{T}\Delta c_p(T)\,dT\,d\beta \\ + \ln\gamma_2(T, x),\end{aligned} \tag{5.5}$$

where $\Delta h^{\mathrm{fus}}(T_m)$ is the enthalpy of fusion of solvent at the normal melting point T_m, $\Delta c_p(T)$ is the difference between the heat capacity between pure solid and liquid solvent, which can be treated in a similar manner to the heat capacity term in explicit models for solubility discussed in Chapter 2, and γ_2 is the solvent activity coefficient in solution (defined with respect to Raoult's Law).

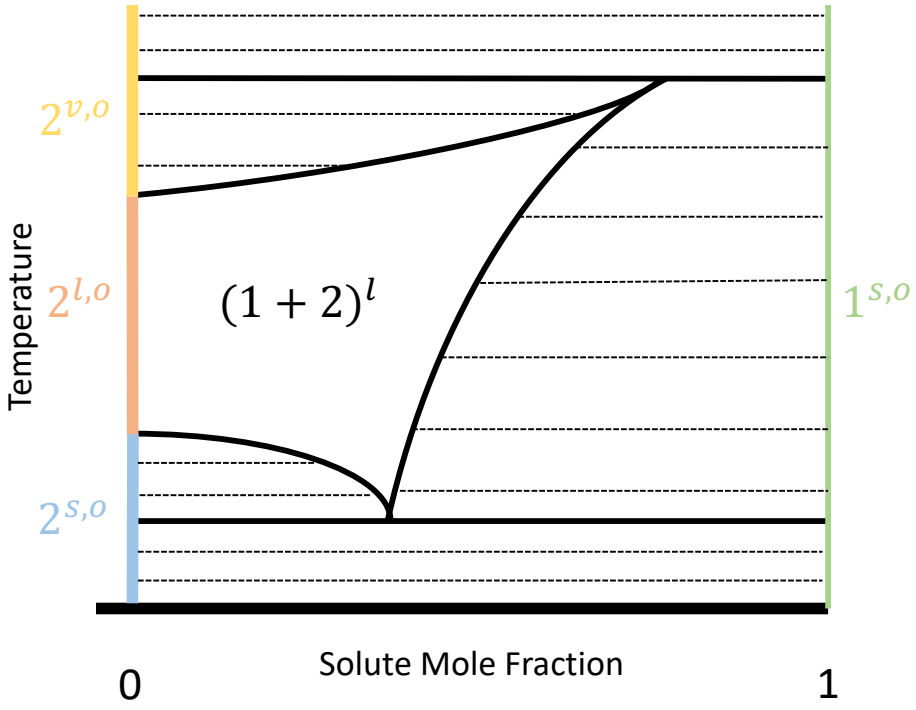


Figure 5.1: Example phase diagram for binary system comprising components 1 (solute) and 1 (solvent) between practically relevant temperatures for solubility estimation. Black solid lines correspond to phase boundaries. Black dashed lines correspond to tie-lines. The blue, orange and yellow portion of the y-axis at 0 component 1 mole fraction (i.e., $2^{s,o}$, $2^{l,o}$ and $2^{v,o}$) indicate the range of temperature the solvent (i.e., component B) exists as a pure solid, liquid and vapor, respectively. The green portion of the y-axis at component 1 mole fraction = 1 (i.e., $1^{s,o}$) corresponds to the range of temperature the solute (i.e., component 1) exists as a pure solid.

Furthermore, it is possible to express the boiling point of a binary solution in a manner similar to Eq. (2.35) and Eq. (5.5), for solubility and freezing point, respectively:

$$\ln\left(1-x_1(T)\right) = \frac{\Delta h_2^{\text{vap.}}(T_b)}{RT_b}\left(\frac{T_b}{T}-1\right) + \int_{\beta(T_b)}^{\beta(T)}\int_{T_b}^{T}\Delta c_p(T)dT\,d\beta + \ln\gamma_2(T,x), \tag{5.6}$$

where $\Delta h_2^{\text{vap.}}(T_b)$ is the enthalpy of vaporisation for pure solvent and T_b is the associated normal boiling temperature, $\Delta c_p(T)$ is the difference between the heat capacity between pure vapor and liquid solvent, and γ_2 is the solvent activity coefficient in solution (defined with respect to Raoult's Law), as before. Equation (2.35), (5.5) and (5.6) connect phase equilibria measurements, melting properties, heat capacity and excess solution properties.

For completeness, in addition to Eq. (5.6), boiling curve data can be included in a manner similar to vapor pressure measurements using modified versions of Raoult's Law where, as long as coefficients in the Antoine equation for the pure solvent are known, the temperature at which a solute-solvent solutions boils can be modelled explicitly via:

$$x_2 = \frac{Py_2}{P_2^{\text{sat.}}(T)\gamma_2}, \tag{5.7}$$

where P is total pressure, y_2 is the mole fraction of solvent in the vapor phase (which is usually always one for the systems encountered in crystallisation applications due to non-volatility of solutes (i.e., the solute has no vapor-pressure)) and $P_2^{\text{sat.}}(T)$ is the vapor pressure of pure solvent at T. It should be noted that this assumes the vapor phase behaves like an ideal gas.

The aforementioned relationships, coupled with an excess Gibbs free energy model such that γ_1 and γ_2 can be evaluated, make it possible to develop thermodynamically consistent descriptions for solute-solvent systems and has been applied in the literature. For example, consider results presented in [145], for the sucrose-water system. In this work, primary focus is the development of a model describing the activity of water in sucrose-water solutions where the model structure and parameter values of which are informed by simultaneously correlating phase equilibria and thermodynamic properties, all of which relate in some manner to the parameters in Eq. (2.35), (5.5) and (5.6). For the sucrose-water liquid mixture, the authors use an four-suffix Margules-type equation to describe the excess Gibbs free energy coupled with an ideal contribution defined relative to Raoult's Law. Phase equilibria are incorporated using Eq. (2.35) for solubility data, Eq. (5.5) for freezing point data and Eq. (5.7) for boiling point data.

A similar application is described in [80] (cited extensively in Chapter 3). Here, thermodynamic theory is coupled with mathematical modeling to develop 'new equations describing the thermodynamic properties of the glycine + H_2O system [...] from previously published thermodynamic data' and the 'primary aim of this work is to develop a thermodynamic equation of state for glycine(aq) capable of describing key measurements of the system's properties within their

experimental uncertainty'. The author constructs explicit models describing the Gibbs free energy of glycine-water liquid mixtures as a function of temperature and composition, as well as the pure α-glycine crystal as a function of temperature. For the glycine-water liquid mixture, the author uses a Pitzer-like equation for non electrolytes as an excess Gibbs free energy contribution defined relative to a hypothetical ideal one molar glycine-water mixture. For the pure glycine crystal, it's heat capacity is modelled using a cubic spline, allowing calculation of the partial molar Gibbs free energy of glycine in the ideal hypothetical solution as a function of temperature, and inclusion of solubility data during model parameterisation.

In both works, all (or subsets of) available experimental data reported in the literature for the system under study are collated, before simultaneously critically assessed and modelled, yielding what is considered a consistent set of parameterised models for the Gibbs free energy of the various phases that adequately describe reviewed data, from which all other thermodynamic properties may be derived, in principle. The types of data used to inform the structure and parameter values of the models used cover heat capacity (single and multi-component phases) and enthalpy changes (upon dilution, dissolution and mixing), as well as phase equilibrium (i.e., vapor-liquid equilibrium in the form of vapor pressure, isopiestic molality, relative humidity, or boiling point measurements, as well as solid-liquid equilibria in the form of temperature dependent solubility and freezing point measurements).

The aforementioned examples illustrate the application of general thermodynamic theory to the problem of simultaneously describing multiple types of experimental thermodynamic data using mathematical models. From a thermodynamic perspective (as discussed in Chapter 2), it is also possible to estimate the phase diagram by finding the distribution of components that minimises the system Gibbs free energy (with a fixed total amount of component A and B) at fixed temperature and pressure where the system Gibbs free energy can be conveniently described as sum of the Gibbs free energy of each phase:

$$\beta G^{\text{sys.}} = \beta G_1^{s,\circ}(T) + \beta G_2^{s,\circ}(T) + \beta G_2^{l,\circ} + \beta G_2^{v,\circ} + \beta G_{1,2}^{l,mix}(T,x) \tag{5.8}$$

where βG^{α} is the extensive Gibbs free energy of phase α and β retains the same definition as previously specified. As an aside, it should also be noted that, for the case where the solid form of a solute is known in addition to the desired solvent for which the solubility of said solute is required, it is possible to estimate solubility by finding the distribution of components that minimises the Gibbs free energy (with a fixed total amount of solvent and solute) of the sub-system comprising pure crystal and solute-solvent liquid mixture only, giving:

$$\beta G^{\text{sys.}} = \beta G_1^{s,\circ}(T) + \beta G_{1,2}^{l,mix}(T,x). \tag{5.9}$$

To estimate a phase diagram or, as in the case required in the thesis, a single phase boundary, functions describing the Gibbs free energy of all relevant phases are required. In a manner similar to

the chemical potential, the Gibbs free energy is not directly measureable; however, it is known (c.f. Chapter 3 and 4) that many experimentally measureable quantities can be related to the underlying Gibbs free energy functions required to estimate any phase boundary in a thermodynamically consistent manner.

5.2.1 The CALPHAD Method

In addition to, but not entirely distinct from, ad-hoc applications, the simultaneous modelling of thermodynamic data and phase equilibria has been formalised in the so-called CALPHAD (**CAL**culation of **PHA**se **D**iagrams) method. The CALPHAD method was developed with a particular focus on using experimental data, mathematical modelling and equilibrium thermodynamics to estimate phase diagrams of often complex multi-phase, multi-component alloy systems and interested readers should refer to [199, 200] for a historical perspective on it's development. Before discussing the approach in any detail, many examples of CALPHAD assessments can be found in the 'CALPHAD Computer Coupling of Phase Diagrams and Thermochemistry' journal and interested readers looking for practical applications should refer there. Further, as mentioned in [201], there are various texts covering the CALPHAD method in great detail and interested readers should refer to [198], [202], and [203] for comprehensive treatments, as only a brief discussion is presented here.

The CALPHAD method can be interpreted as a systematic, iterative procedure often referred to as an 'assessment' and comprises two stages: critical analysis and optimisation. In the critical analysis stage, relevant experimental and theoretical data for the system under study is collected from the available sources (i.e., literature, databases, etc). During the collection of data, it is recommended that all data sources are reviewed for the 'appropriateness of experimental methods, systematic uncertainties, impurities, etc', while an initial attempt to define the 'number of independent model parameters for each phase, consistent with quantity and quality of available information' should be made [198]. Assuming the initial models for all relevant phases are correct as well as the assessed experimental data, the next step is to estimate the parameters in the chosen model by finding the parameter values that allow models to describe available data. This is usually approached as a least-squares optimisation [198]; however, recent works have explored the use of Bayesian techniques [204], with the goal of allowing uncertainty propagation from experimental data and quantification in model predictions. Usually, this process is repeated by varying model structure and parameter constraints, as well as weighting and/or inclusion of data.

In a CALPHAD assessment, the Gibbs free energy of all phases are modelled explicitly and, as mentioned already, experimental data is used to identify the structure of the underlying Gibbs free energy models, as well as the values of parameters they comprise. Perhaps most simply, the Gibbs free energy of a phase with fixed composition is usually expressed as a power series in

temperature [198]:

$$G = a_0 + a_1 T + a_2 T \ln T + a_3 T^2 + a_4 T^{-1} + \ldots \tag{5.10}$$

where $a_1, a_2, \ldots$ are free parameters found by correlating experimental data. Equation. (5.10) is usually only appropriate over limited temperature ranges and may lead to significant errors if used for extrapolation [198]. Crucially for stable components, the heat capacity of which is experimentally measurable, it is possible to evaluate an appropriate model structure and some parameter values given that for Eq. (5.10):

$$c_p = \frac{\partial^2 \beta G}{\partial \beta \partial T} = -a_2 - 2a_3 T - 2a_4 T^{-2} - 6a_5 T^2 + \ldots. \tag{5.11}$$

The extensive Gibbs free energy of a phase with a composition dependence is usually interpreted through different contributions:

$$G_\alpha = G_\alpha^{\text{srf}} + G_\alpha^{\text{phys}} - T S_\alpha^{\text{cnf}} + G_\alpha^{\text{E}}, \tag{5.12}$$

where G_α^{srf} corresponds to a contribution to the Gibbs free energy associated to a 'surface of reference', G_α^{phys} corresponds to a contribution to the Gibbs free energy associated to a physical model describing a particular phenomenon (e.g., magnetic transitions), G_α^{E} is the excess Gibbs free energy and S_α^{cnf} is the configurational entropy [198].

For the systems discussed in this thesis, the only phase with a composition dependence is the solute-solvent liquid mixture. In CALPHAD terms, this is referred to as a 'substitutional-solution' and is usually modelled without any 'physical' contribution in which case Eq. (5.12) becomes:

$$G_\alpha = \sum_i^n n_i g_i^{\text{l},\circ} + RT \sum_i^n n_i \ln x_i + G_\alpha^{\text{E}}, \tag{5.13}$$

where the excess Gibbs free energy contribution is modelled using the Redlich-Kister power series (c.f. Chapter 4), expressed for the molar excess Gibbs free energy via:

$$g_\alpha^{\text{E}} = \sum_i^{n-1} \sum_j^n x_i x_j L_{ij}, \tag{5.14}$$

where $L_{ij} = \sum_\nu^k (x_i - x_j)^\nu \, {}^\nu L_{ij}$. The appropriate number of coefficients required for a given phase cannot be known *a priori* and is therefore determined subjectively during the optimisation stage based on the quality, type and distribution of available experimental data. When the experimental data used during an assessment allows, the term ${}^\nu L_{ij}$ can be interpretted as a function of temperature and usually in the form of a power series in temperature[1] as for phases with fixed composition which, for six coefficients is expressed as [205]:

$${}^\nu L_{ij} = {}^\nu a_{ij} + {}^\nu b_{ij} T + {}^\nu c_{ij} T \ln T + {}^\nu d_{ij} T^2 + {}^\nu e_{ij} T^3 + {}^\nu f_{ij} T^{-1} \tag{5.15}$$

[1] with some additional terms for fitting convenience

where, again, the appropriate number of coefficients required to describe a given phase is determined during the optimisation stage of an assessment. Assessments with temperature dependent interaction parameters including all six parameters appear to be limited, and it is common to use only the first three [205, 198]. In [205], it was suggested that interaction parameters could instead be described by a two parameter exponential equation to avoid observed high temperature artifacts occasionally observed with the extrapolation of power series expansions to high temperatures, giving:

$$ {}^{v}L_{ij} = {}^{v}h_{ij} \exp\left(-\frac{T}{{}^{v}\tau_{ij}}\right), \tag{5.16} $$

where ${}^{v}h_{ij}$ and ${}^{v}\tau_{ij}$ are flexible parameters and ${}^{v}\tau_{ij} > 0$. Interested readers are also directed to more recent works [206, 207] by the same author in [205], where the exponential temperature dependence is discussed in greater detail, as well as that described in [208].

As an aside, it should be noted that in addition to the popular Redlich-Kister excess Gibbs free energy model cited in seminal texts on the CALPHAD method, as well as being used in the majority of CALPHAD assessments, the standard UNIQUAC equation [43] was recently implemented [209] in an open-source Python package (i.e., OpenCalphad [210]) developed for completing CALPHAD assessments.

A CALPHAD assessment is, generally speaking, analogous to the processes described in the ad-hoc examples cited previously [80, 145], aside from subtle choices around modelling approach and property calculation. In both approaches, the fundamental idea is to use experimentally measurable thermodynamic quantities and phase equilibria to generate thermodynamically consistent mathematical representations of systems such that any property of the system can be reliably estimated.

Typically, the development of Gibbs free energy models through simultaneous representation of experimental data rely solely on the data available at the time of the study and, as such, the quality of resulting models is limited by this. As shown in both ad-hoc applications and those based on the CALPHAD approach, it is common practice to include directly measured temperature dependent phase equilibrium data (e.g., solubility) during model parameterisation, regardless of the intended application. However, in line with the focus of this thesis, there have been no attempts to understand how this general idea of constructing thermodynamically consistent Gibbs free energy models from experimental data can be used to estimate temperature dependent solubility, specifically. Therefore, a key question is can this general approach can be used to estimate solubility in regions where it hasn't yet been measured.

5.2.2 The Relationship Between CALPHAD-like Gibbs free Energy Models and Experiments

In principle, developing a model for the Gibbs free energy of an arbitrary phase in a system can be approached using any appropriate modelling technique (e.g., power series, cubic spline, neural network, physical model, etc.). Given the efficacy of models used during CALPHAD assessments (i.e., Eq. (5.10) and (5.12)) as illustrated by the volume of assessments reported in the literature, these are used as a basis for the remainder of this chapter. As mentioned already, the Gibbs free energy of a pure crystalline solute and solute-solvent liquid mixture are required to estimate solubility either through minimisation of the composite Gibbs free energy (or solution of Eq. (2.31)). In this section, the relationship between experimentally measureable thermodynamic quantities and the models used to describe both phases are discussed.

Assuming, the Gibbs free energy of pure crystalline solute can be expressed using a CALPHAD power series:

$$\beta g_i^{\mathrm{s},\circ}(T) = a_0\prime + a_1\prime\left(\frac{T_0}{T} - 1\right) + f(T). \tag{5.17}$$

where $f(T)$ is an unknown function of temperature which can, over a given range of temperature, be completely specified by knowledge of the pure crystal solute heat capacity, assuming this is experimentally accessible. Furthermore, if heat capacity data for the pure crystal is available from 0 K, it is possible to estimate a_0 and a_1, as illustrated in Section 5.6. From a mathematical perspective the parameters in Eq. (5.17) are simply free parameters, an appropriate value for which can be found by fitting experimental data. However, from a thermodynamic perspective, it is possible to relate the Gibbs free energy of a phase with constant composition to enthalpy and heat capacity:

$$\beta g_i^{\mathrm{s},\circ}(T) = \beta_0 g_i^{\mathrm{l},\circ}(T_0) + \beta_0 h_i^{\mathrm{l},\circ}(T_0)\left(\frac{\beta}{\beta_0} - 1\right) + \int_{\beta_0}^{\beta}\int_{T_0}^{T} c_{p,i}^{\mathrm{l},\circ}(T)\,dT\,d\beta, \tag{5.18}$$

giving, $a_0\prime = \beta_0 g_i^{\mathrm{s},\circ}(T_0)$, $a_1\prime = \beta_0 h_i^{\mathrm{s},\circ}(T_0)$, and $f(T) = \int_{\beta_0}^{\beta}\int_{T_0}^{T} c_p(T)\,dT\,d\beta$.

The solute-solvent liquid mixture can be described using Eq. (5.13). The form of this equation is identical to that of a solution interpreted in terms of an ideal solution contribution (in terms of Raoult's Law) and an excess contribution. Following the standard CALPHAD approach, the temperature dependent function associated to the pure liquid components is given by Eq. (5.17). For simplicity, no specific structure is specified for the excess Gibbs free energy contribution, beyond the fact it is a function of temperature and composition.

For a set of experimentally measureable quantities, it is possible to enumerate how they relate to parameters in each model. Common measurements reported in literature and used during development of thermodynamic models (based on work discussed in Chapter 3 and 4) for systems discussed in this thesis are presented in Table 5.1. It should be noted that pure solute melting

properties (i.e., enthalpy of fusion and normal melting temperature) have been omitted intentionally (c.f. Chapter 2), in keeping with the general aim of this thesis.

Table 5.1: Model parameter–data relationships, where x_1 denotes solubility, a_2 denotes solvent activity in solution, $\Delta H^{\text{dil.}}$ denotes enthalpy of dilution, $\Delta H^{\text{sol.}}$ denotes enthalpy of solution and c_p^{soln} denotes solution heat capacity. † denotes property must be measured at temperatures other than T_0. ‡ denotes single estimate of property required if measured at T_0.

Parameter	x_1	a_2	$\Delta H^{\text{dil.}}$	$\Delta H^{\text{sol.}}$	c_p^{soln}
$g_1^{\text{l},\circ}(T_0)$	✓‡				
$h_1^{\text{l},\circ}(T_0)$	✓†			✓‡	
$c_{\text{p},1}(T)$	✓			✓	✓
βG^E	✓	✓	✓	✓	✓

Before discussing the relationship between experiments and parameters described in Table 5.1, it should be noted that it is possible to evaluate parameters associated to the pure liquid solvent in a manner identical to that described previously for the pure crystalline solute, in principle, given that it is a stable, experimentally observed phase. In contrast, from a practical perspective, all parameters associated to the pure hypothetical liquid solute must be found indirectly through a specific set of experimentally measurable quantities, given that pure liquid solutes at temperatures below their assumed normal melting point are unstable and therefore rarely observed using conventional experimental techniques (c.f. Chapter 2). The contribution associated to the hypothetical pure liquid solute are discussed in greater detail next.

From the experiments listed in Table 5.1, there is only one measurement that gives information about the value of $g_1^{\text{l},\circ}(T_0)$ and that is solubility. In theory, the experimental conditions at which solubility measurements are required to estimate $g_1^{\text{l},\circ}(T_0)$ is reduced to one if the solubility is measured at T_0. However, this comes with the caveat that a reliable model for $\beta G^E(T, x)$ is known at T_0 as a function of solution composition. Further more, $h_1^{\text{l},\circ}(T_0)$, $c_{\text{p},1}^{\text{l},\circ}(T)$ and $\beta G^E(T, x)$ are required if solubility data at temperatures other than T_0 are used. The key finding from this is that, for this collection of experiments and the modelling framework described, at least one solubility measurement is required such that all parameters are evaluated, if solubility is to be estimated at other temperatures.

There are only two types of experimental measurement that give information about the value of $h_1^{\text{l},\circ}(T_0)$ and they are solubility (at temperatures other than T_0) and enthalpy of solution. Like with solubility and $g_1^{\text{l},\circ}(T_0)$, the number of experimental state points at which the enthalpy of solution is required to estimate $h_1^{\text{l},\circ}(T_0)$ is reduced to one if it is measured at T_0 and an arbitrary solution concentration (below saturation). However, like with solubility, this comes with the caveat that a

reliable model for $\beta G^E(T,x)$ is known at T_0 and as a function of solution composition. Again, if the enthalpy of solution is not at T_0, then both $c_{\mathrm{p},1}^{\mathrm{l},\circ}(T)$ and $\beta G^E(T,x)$ are required.

Finally, there are three types of experiments that give information about $c_{\mathrm{p},1}(T)$ and that is solubility, enthalpy of solution, and solution heat capacity. If, for example, the functional form of the heat capacity is assumed to be that of an expansion-type (e.g., $c_{p,1}^{\mathrm{l},\circ}(T) = c_{\mathrm{p},1}^{\mathrm{l},\circ}(T_0) + \ldots$) then it is possible to estimate the first term (i.e., $c_{\mathrm{p},1}^{\mathrm{l},\circ}(T_0)$) with the solution heat capacity at T_0. Again, this comes with the caveat that a reliable model for $\beta G^E(T,x)$ is known at T_0 as a function of solution composition.

In theory, it is possible to constrain $\beta G^E(T,x)$ with an additional set of experiments that do not include $g_1^{\mathrm{l},\circ}(T_0)$, $h_1^{\mathrm{l},\circ}(T_0)$ or $c_{\mathrm{p},1}^{\mathrm{l},\circ}(T_0)$. These include measurements based on vapor liquid equilibrium (i.e., vapor pressure, isopiestic molality, relative humidity, etc), the enthalpy of dilution (which takes numerous forms), and potentially even light scattering [211]. Additionally, it is possible to use other phase boundaries that enclose the liquid solution (i.e., boiling curve and freezing point depression), if the Gibbs free energy of the adjacent phase (i.e., pure vapor and pure crystalline solvent, respectively) is known. All of these examples yield information about different partial derivatives of the $\beta G^E(T,x)$ surface.

From a practical perspective, it is possible to frame the construction of a Gibbs free energy surface for a binary liquid solution as a two step process where certain experimental data is used to independently construct $\beta G^E(T,x)$ and then evaluate the remaining parameters associated to the sub-cooled liquid solute. Alternatively, it it possible to view the processes as a single step, where all data is used simultaneously. Regardless, the primary goal is to find functional forms for $c_{\mathrm{p},1}^{\mathrm{l},\circ}(T)$ and $\beta G^E(T,x)$ that describe the available experimental data with the hope that the resulting surface leads to reliable solubility estimates.

Assuming complete determination of Gibbs free energy functions of stable, experimentally observable pure solvent phases (i.e., crystalline, liquid and vapor), as well as pure crystalline solute phases is reasonable given our ability to measure the heat capacity of stable compounds. This would be analogous to the unary database developed as part of the CALPHAD modelling concept, as originally published by [212]. Treating the determination of model structure and parameter values that are directly related to the temperature dependence of pure stable component heat capacity has practical advantages. First, it reduces the number of parameters that would be required during the optimisation stage which may reduce the complexity of the objective function and lead to more robust results with reduced computational effort, depending on the algorithm used. Secondly, this eliminates data types from the global optimisation step that would otherwise contribute to the chosen objective function. If framed as a single objective optimisation, this could introduce unnecessary complexity and poor model parameterisation, as discussed in a recent work assessing the performance single and multi-objective optimisation for excess Gibbs free energy

models [213].

In summary, for the experimentally measureable quantities in Table 5.1, it is possible to construct combinations of experiments that allow parameterisation of the Gibbs free energy models required to estimate solubility as a function of temperature. Most importantly, in the broader context of the work presented in this thesis, it is possible to do so with only a single estimate of solubility, retaining a key property of the approaches described in Chapter 3 and 4. In the remainder of the chapter, the applicability of this result is analysed on 'real' experimental data for the sucrose-water system. More specifically, available experimental data is collated and modelled for the purpose of understanding how well the associated solubility curve can be represented. For the purpose of this chapter, modelling is approached using the aforementioned Gibbs free energy models frequently used in the CALPHAD method.

5.3 Computational Details

For the purpose of the work presented in this chapter, a series of Python modules were developed to allow simple calculation of thermodynamic properties from parameterised Gibbs free energy models for single and multi-phase system comprising either single or multiple components, as well as the parameterisation of Gibbs free energy modes by 'fitting' to experimental data. It should be noted that the development of a comprehensive software package for distribution as either open-source or commercial use is beyond the scope of this thesis.

5.3.1 Automatic Differentiation

'Automatic Differentiation (AD)' is a computational technique used to evaluate partial derivatives of mathematical function specified in terms of computer code. As noted in [214], in contrast to alternative computational differentiation techniques, AD voids the need to express mathematical functions in 'expression strings' (c.f. symbolic differentiation) and eliminates approximation error (c.f. numerical differentiation). Given that any thermodynamic quantity can be expressed as a partial derivative (or combination of) of the underlying Gibbs free energy (or relevant thermodynamic potential), Automatic Differentiation was used to calculate thermodynamic properties from underlying Gibbs free energy models in Chapter 3 and 4, and plays a crucial role in the calculations described in this chapter.

5.3.2 Python Code

For calculation of thermodynamic quantities from pre-assigned Gibbs free energy models or parameterisation of models based on experimental data, the following three core modules were developed:

- Models.py
- DataFunctions.py
- DataProcesses.py,

all of which are discussed individually.

Models.py

In Models.py, Gibbs free energy models required for subsequent calculation of derived thermodynamic quantities, parameterisation by fitting to experimental data, or both, are stored. Although it may have been reasonable to apply an 'Object Orientated Programming (OOP)' philosophy and interpret a Gibbs free energy model as 'class', it was decided to encapsulate a Gibbs free energy model in a Python function (more akin to a Functional Programming philosophy). This choice was motivated by the manner in which the AutoGrad Python library allows calculation of derivatives and the choice is not expected to have any material impact on performance in the context of the computations described herein.

The structure of a Gibbs free energy Python function in the Model.py module is as follows:

```
def GibbsFreeEnergyModel(paramater_dict, *state_data):

    # Unpack state data
    B = state_data[0]
    P = state_data[1]
    n_list = state_data[2:]

    # Calculate BG
    (insert calculation here)

    return BG
```

where it should be noted that that the function takes two arguments, the first being parameter_dict which is a Python dictionary that holds all parameters (e.g., binary interaction paramaters in the UNIFAC model). The second, *state_data, denotes a Python iterable (e.g., tuple, list, etc.), for all state properties, including: $\beta = 1/(RT)$, pressure and the number of moles of each component in the model (which may be one or more). Throughout this work, the first and second values in *state_data are always designated as β and pressure, respectively; however, this is not necessary, as long as the un-packing process is consistent. It should be noted that the property calculated by Python function is βG.

In principle, with Gibbs free energy models stored as Python functions in Models.py, users can derive thermodynamic quantities using AutoGrad. However, the additional Python modules DataFunctions.py and DataProcesses.py have been developed to store re-usable code snippets frequently encountered when calculating thermodynamic properties of interest and fitting experimental data.

DataFunctions.py

The DataFunctions.py module acts as a placeholder to store user-defined Python functions that manipulate Gibbs free energy models (defined in Models.py) and generate thermodynamic quantities (e.g., osmotic coefficient, differential enthalpy of dilution, solubility, etc), typically using AutoGrad. The basic structure of a DataFunction is as follows:

```
def DataFunction(model_dict, param_dict_dict, *state_data):

    # Calculate thermodynamic property of interest

    return Property
```

where it should be noted that a DataFunction will take three arguments, the first being model_dict, which is a Python dictionary that collects all Gibbs free energy models required to calculate the property of interest (e.g., pure solid and liquid mixture for solubility), the second being param_dict_dict which is another Python dictionary used to collect the Python dictionaries associated to each Gibbs free energy model in model_dict, and the final argument is *state_data which, in most cases takes a similar form to *state_data used by Python functions defined in Models.py; however, can be defined in any manner that a user deems appropriate to facilitate calculation of the thermodynamic property of interest.

DataProcesses.py

The DataProcesses.py module acts as a placeholder to store Python functions that return a user specified 'goodness of fit' measure between thermodynamic quantities of interest derived from Gibbs free energy models and corresponding set of experimental data. The purpose of this is to simplify the process of fitting Gibbs free energy models to experimental data, as the calculated goodness of fit value can be fed into a user-defined objective function. For the purpose of this work, this was typically chosen as the 'sum of errors squared', which contributed to an objective function used by a numerical optimisation algorithm taken from NumPy [215]. The basic structure of a DataProcess is as follows:

```
def DataProcess(df, function, model_dict, param_dict_dict):

    # Calculate goodness of fit to experimental data
    # based on current parameters values stored in
    # param_dict_dict

    return Goodness_of_fit
```

where a DataProcess Python function takes four inputs, the first being df, which denotes a Pandas [216, 217] dataframe Python object that holds the experimental data that the user wishes to compare a set of parameterised Gibbs free energy models to, function is a DataFunction defined for the thermodynamic property of interest, model_dict_dict is a Python dictionary containing the Python dictionaries containing parameter objects for each of the Gibbs free energy model Python functions (defined in Models.py) required for the analysis and param_dict is a Python dictionary holding the parameters associated to each Gibbs free energy model.

It should be noted that, in the current form, a DataProcess must be written for all types of experimental data that may be used during a fitting process. More specifically, is a user wants to fit a both vapor pressure data (reported in terms of vapor pressure, temperature and solution composition) and solution heat capacity data (reported in terms of solution heat capacity, temperature, pressure and solution composition), the user would have to generate and concatenate the outputs to generate the desired objective function.

5.4 The Sucrose-Water System

In the previous section, it was shown how different types of experimental measurements inform an appropriate model structure and associated parameter values when developing Gibbs free energy models for solubility prediction. As mentioned briefly in Chapter 4, there appears to be few systems with the characteristics suitable for crystallisation (i.e., at least one component stable in the liquid state (solvent) and another stable in the solid state (solute)), where phase equilibria have been accurately determined across all relevant temperatures and atmospheric pressure, as well as complete characterisation of single and multi-component phases, in terms of the experimental measurement techniques detailed in the previous section; however, one possible exception is sucrose-water. As illustrated in Fig. 5.2, phase boundaries of the sucrose-water system are well defined through direct experimental measurements.

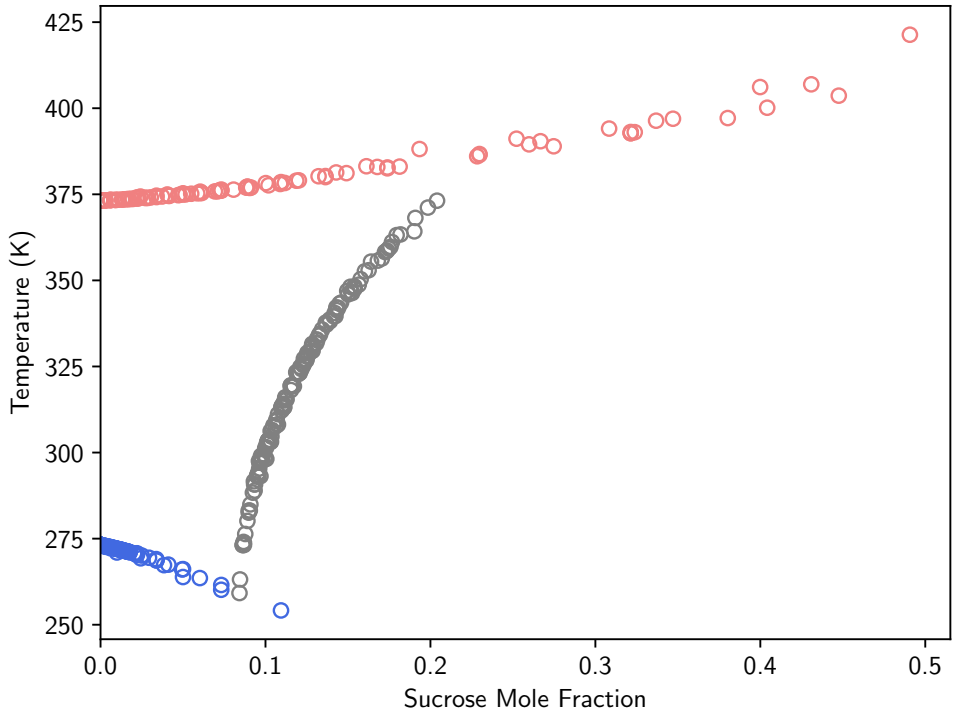


Figure 5.2: Sucrose-water phase equilibria at 1 atm. Open circles correspond to direct measurements reported in the literature. Red denotes boiling point measurements [218, 219, 220, 221, 222, 223]. Grey denotes solubility measurements [146, 147, 148, 149, 150, 151]. Blue denotes freezing point measurements [181, 182, 148, 150, 183, 184].

For completeness, the work described in this chapter related to the sucrose-water is focused on the anhydrous form of sucrose only, noting that several sucrose hydrate forms have been reported in the literature [148]. Regardless, in addition to well defined phase boundaries, many studies reporting thermodynamic properties of the sucrose-water solution phase are present in the literature, as shown by an extensive data compilation discussed in [145].

In addition to the aforementioned work described in [145], modelling thermodynamic properties and phase equilibria of the sucrose-water system has been discussed by several authors (tabulated in Table 5.2).

Table 5.2: Sucrose-water thermodynamic modelling studies

Author	Year	Citation
Abed et al.	1992	[224]
Abderafi et al.	1994	[225]
Catte et al.	1994	[226]
Catte et al.	1995	[227]
Peres and Macedo	1996	[228]
Peres and Macedo	1997	[11]
Kuramochi et al.	1997	[107]
Jonsdottir and Rasmussen	1999	[57]
Spiliotis and Tassios	2000	[49]
Ferreira et al.	2003	[229]
Tsavas et al.	2004	[230]
Perozin et al.	2007	[231]
Nowak et al.	2009	[232]
Van der Sman	2017	[233]
Paese et al.	2020	[234]

Most studies listed in Table 5.2 have been discussed in [145] and [234], respectively. As noted by these authors, many sucrose-water thermodynamic modelling studies have focussed on the determination of appropriate group contribution functional groups and/or associated group contribution parameter values for group contribution excess Gibbs free energy models. Examples of this type of work include [224] where interaction parameters between molecular groups not available in published tables (at the time of original publishing) for the original UNIFAC model were evaluated by correlating temperature dependent solubility data for binary and ternary sugar mixtures, [227] where a new expression based on the UNIFAC equation that includes an association/solvation term to capture behavior thought to occur in sugar-water solutions, [49] where the so-called 'S-UNIFAC model was proposed, providing new groups specifically for modelling liquid mixtures containing sugars was proposed and their interaction parameters against experimental data (which was later extended by [230]) and [229] where the so-called 'A-UNIFAC' model was proposed and correlated against a large selection of data giving good correlations and applications to mixed systems.

Aside from works investigating the application of group contribution excess Gibbs free energy models to the sucrose-water system (again, as pointed out in [145] and [234]), and perhaps more relevent to the work described in the remainder of this chapter, several authors have as-

sessed the performance of correlative excess Gibbs free energy models in the context of describing experimental sucorse-water thermodynamic data. Examples of this type of work include [225] where parameters in the Peng-Robinson and Lee-Kesler equations of state, as well as the NRTL excess Gibbs free energy model were correlated to boiling point measurements at atmospheric pressure for binary and ternary sugar-water mixtures (including sucrose-water) published in the same work, [226] where various types of experimental thermodynamic data (including solubility, freezing point, boiling point, solvent activity, enthalpy of dilution and solution heat capacity) were correlated using using a modified UNIQUAC equation for the excess Gibbs free energy, [228] where phase equilibria (i.e., solubility, freezing point and boiling point) and solvent activity data were correlated using the UNIQUAC equation with temperature dependent interaction parameters, [231] the PRSV equation of state was used to correlate sucrose-water solvent activity data and then used to predict boiling temperatures in good agreement with direct measurements, [233] the Flory-Huggins excess Gibbs free energy model was used to correlate temperature dependent solubility data, and the often cited (in this thesis) [145] where the biggest collection of thermodynamic data (including solubility, boiling and freezing point, solvent activity, enthalpy of dilution and solution heat capacity) is correlated with the goal of producing an excess Gibbs free energy model describing the activity of water in sucrose-water liquid mixtures.

In summary, the aforementioned works have mostly focussed on the problem of using available experimental data to determine appropriate model structures, and/or parameterise a predetermined model—trying to 'predict' solubility from all other available experimental data has not been addressed. Furthermore, with the exception of [145], many studies have sampled only a subset of available data. Finally, there are no examples where a Redlich-Kister polynomial excess Gibbs free energy model has been used to describe the sucrose-water liquid mixture. All of the above are addressed, in-part, in the remainder of this chapter.

5.5 Data Review

The sucrose-water system has been studied through measurement of several different thermodynamic properties, as indicated by a compilation published [145], which cites a vast collection of reports of phase equilibria (i.e., vapor-liquid and solid-liquid) as well as thermochemical (i.e., activity, enthalpy and heat capacity) data. The authors refer to an earlier compilation [235], as well as previous work of the lead author [236], as basis for their compilation.

In this chapter, the compilation published in [145] was used as a basis to construct reference data for which a sucrose-water model could be developed. In addition to reports cited in [145], the DETHERM database provided by the Physical Sciences Data-science Service (hosted by the University of Southampton) was used to find additional references, as well as the wider literature

(where relevant).

Broadly speaking, data included in this analysis belong to one of four groups:

1. Cited in [145] and available through a primary source [2]
2. Not cited in [145], available in DETHERM and available through a primary source
3. Cited in [145] and available through DETHERM but not available as a primary source
4. Not cited in [145], not available through DETHERM but available as a primary source

It should be noted that some reports cited in [145] were not included in the data collection stage for a variety of reasons (e.g., primary source inaccessible, primary source in language other than English, etc.,); however, these omissions were restricted to data types and regions on the phase diagram where a meaningful amount of data was already available; thus, it is not expected to have substantial impact on the analysis presented here. In subsequent sections, the data collected is summarised.

5.5.1 Phase Equilibria

Direct measurements of sucrose-water phase equilibria including both boiling and freezing points, as well as anhydrous sucrose solubility have been reported in the literature and are illustrated in Fig. 5.2. As shown, each phase boundary appears to be well defined with data spanning the entire range of practical interest from the perspective of crystallisation process design (i.e., complete knowledge of solubility curve from eutectic to vapor-solid-solution triple point).

Focusing on the solubility data, illustrated separately in Fig. 5.3, it can be seen that all data sets included in this analysis are in excellent agreement with each other.

[2] In this context, a primary source is defined as the document where the data was originally reported

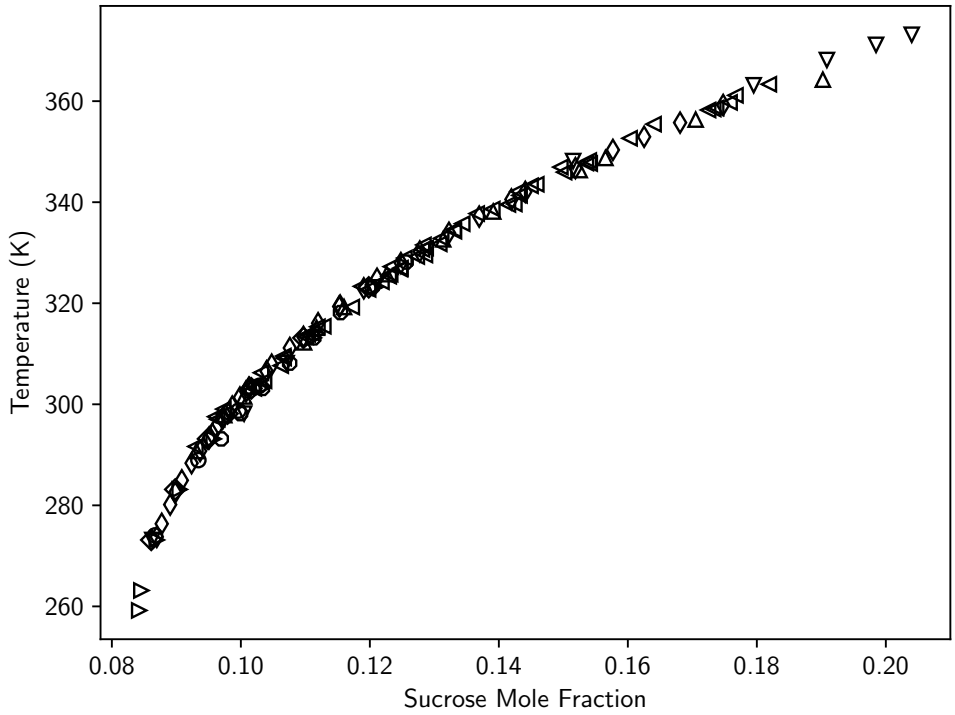


Figure 5.3: Sucrose solubility in water at 1 atm. Open symbols correspond to direct measurements reported in the literature [146, 147, 148, 149, 150, 151].

The lowest temperature at which the solubility has been reported is approximately 260 K in [148] while the highest is approximately 370 K reported in [146]. There appears to be no measurements closer to the assumed vapor-crystal-solution triple point and this could be a direct for future research; however, may present challenges due to high solution concentration, temperatures far from ambient and water volatility.

Measurements of water freezing point temperature in sucrose-water liquid mixtures are presented in Fig. 5.4.

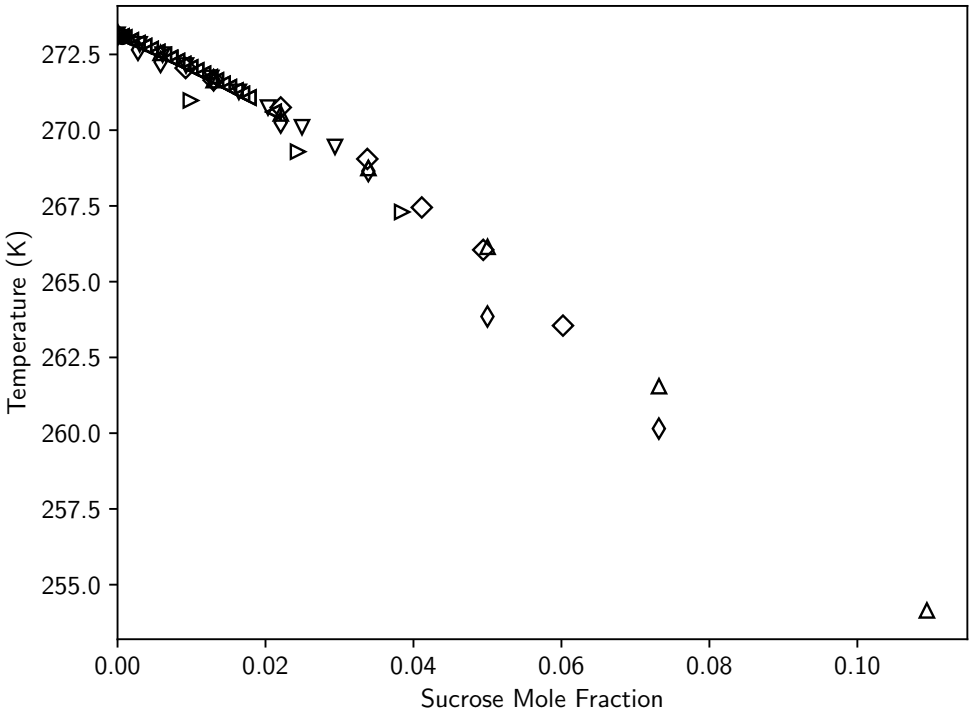


Figure 5.4: Water freezing point temperature in sucrose-water liquid mixtures at 1 atm. Open symbols correspond to direct measurements reported in the literature [181, 182, 148, 150, 183, 184].

As shown, water freezing temperature decreases with increasing sucrose concentration in accordance with the general principle of 'freezing point depression', where the freezing point of a solvent decreases with the addition of solute. Data from independently reported datasets are in fair agreement at solution concentrations up to approximately 0.04 sucrose mole fraction. Three references report data that extend above this towards the assumed eutectic point [148, 184], one of which reports freezing point data beyond the expected eutectic point [148].

Measurements of water boiling point temperature in sucrose-water liquid mixtures are presented in Fig. 5.5.

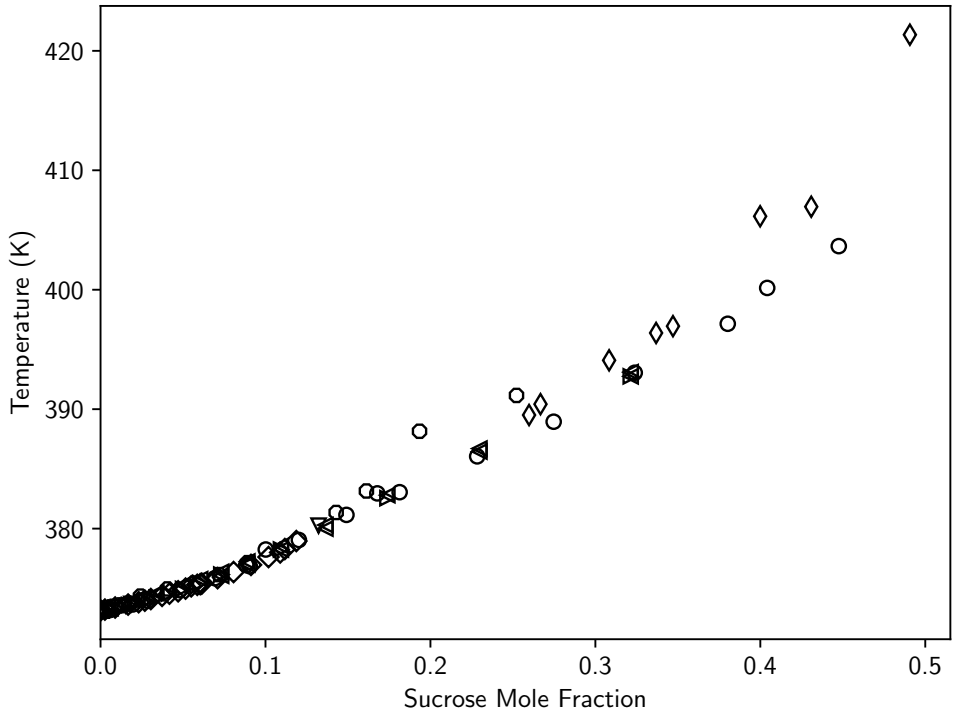


Figure 5.5: Water boiling point temperature in sucrose-water liquid mixtures at 1 atm. Open symbols correspond to direct measurements reported in the literature [218, 219, 220, 221, 222, 223].

As shown, water boiling temperature increases with increasing sucrose concentration in accordance with the general principle of 'boiling point elevation', where the temperature at which a solvent boils increases with the addition of solute. For the references shown, data from independent reports are in good agreement up to 0.2 sucrose mole fraction. Beyond 0.2 sucrose mole fraction, data from independent reports begins to diverge. By comparison to Fig. 5.2, the intersection of the solubility and boiling point curve is likely around 0.2 sucrose mole fraction, which may explain increased variation.

5.5.2 Pure Component Constant Pressure Heat Capacity

For water in both liquid and vapor phases, the 'International Association for the Properties of Water and Steam' (IAWPS) have published a model, referred to in the remainder of this chapter as IAWPS-95, reported to describe the most accurate experimental thermodynamic data within experimental uncertainty, where model development and predictions are presented in [237]. Predictions for the constant pressure heat capacity of both phases derived from the IAWPS-95 model are shown in Fig. 5.6. As shown, the constant pressure heat capacity for water in the liquid state is

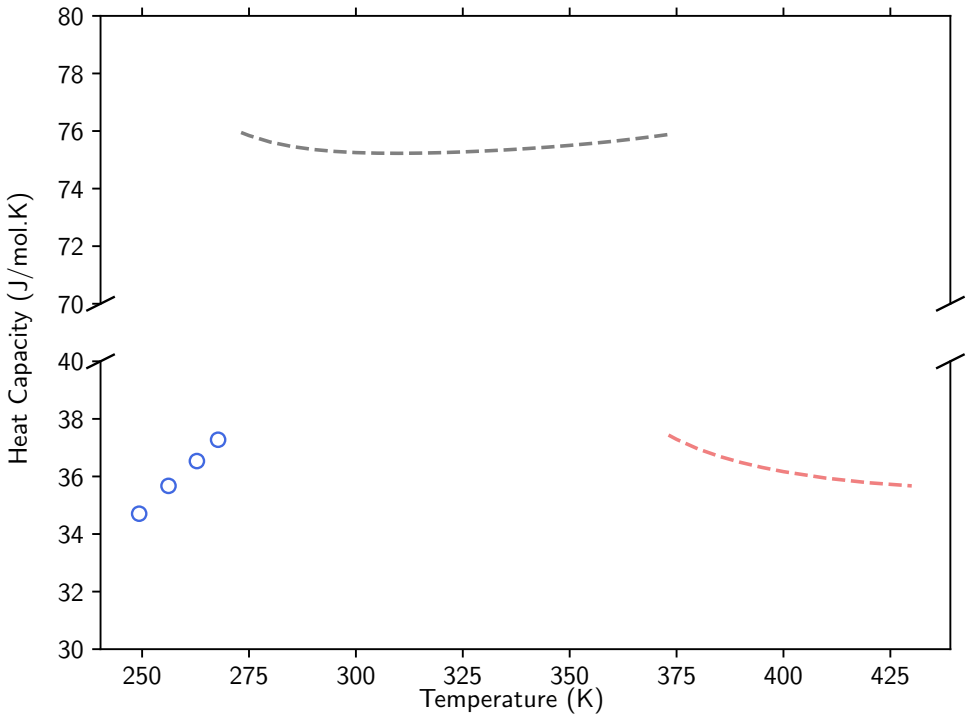


Figure 5.6: Constant pressure molar heat capacity for solid, liquid and vapor water between 260–420 K at 1 atm. Dashed lines correspond to estimations based on IAWPS-95. Open symbols correspond to direct measurements [238]. Blue denotes solid, grey denotes liquid and red denotes vapor.

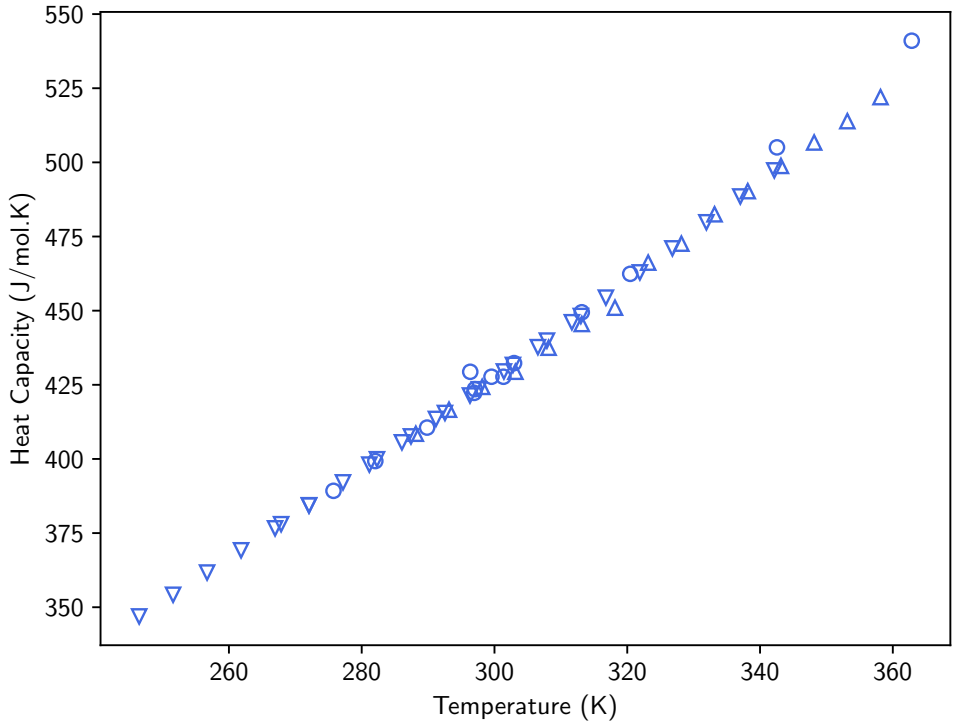


Figure 5.7: Constant pressure molar heat capacity measurements for crystalline anhydrous sucrose between 260–420 K at 1 atm. Dashed lines denote fits. Open symbols denote direct measurements [239, 240, 241].

modelled by a curve starting with a heat capacity of approximately 75.95 J/mol.K at 273.15 K, decreasing to a minimum of 75.23 J/mol.K at 305 K, before increasing to 75.88 J/mol.K at 373.15 K (i.e., normal boiling point). For the vapor phase, the constant pressure heat capacity is modelled by monotonically decreasing curve starting at approximately 37 J/mol.K at 373.15 K and dropping to approximately 35 J/mol.K at 430 K.

For water in the solid phase (i.e., ice), constant pressure heat capacity data was collected from the literature and this data is illustrated in Fig. 5.6, where the data indicates that the constant pressure heat capacity is a linear function of temperature.

Experimental data from independent studies reporting the constant pressure heat capacity of pure crystalline sucrose were collected from the literature, and illustrated in Fig.5.7. As shown, the constant pressure heat capacity increases with increasing temperature. All independent reports shown are in good agreement between 260–420 K and the data appear to have a linear dependence on temperature, in a manner similar to water in the solid state.

5.5.3 Vapor Pressure, Isopiestic Molality and Relative Humidity

The vapor pressure above sucrose-water solutions has been reported by several authors. Briefly, the collection contains measurements from thirteen independent sources from as early as 1910 to as recent as 2018, while numerous apparatus have been described. All measurements are presented in Fig. 5.8, while data in the temperature range is 273.15–323.15 K is shown in Fig. 5.9.

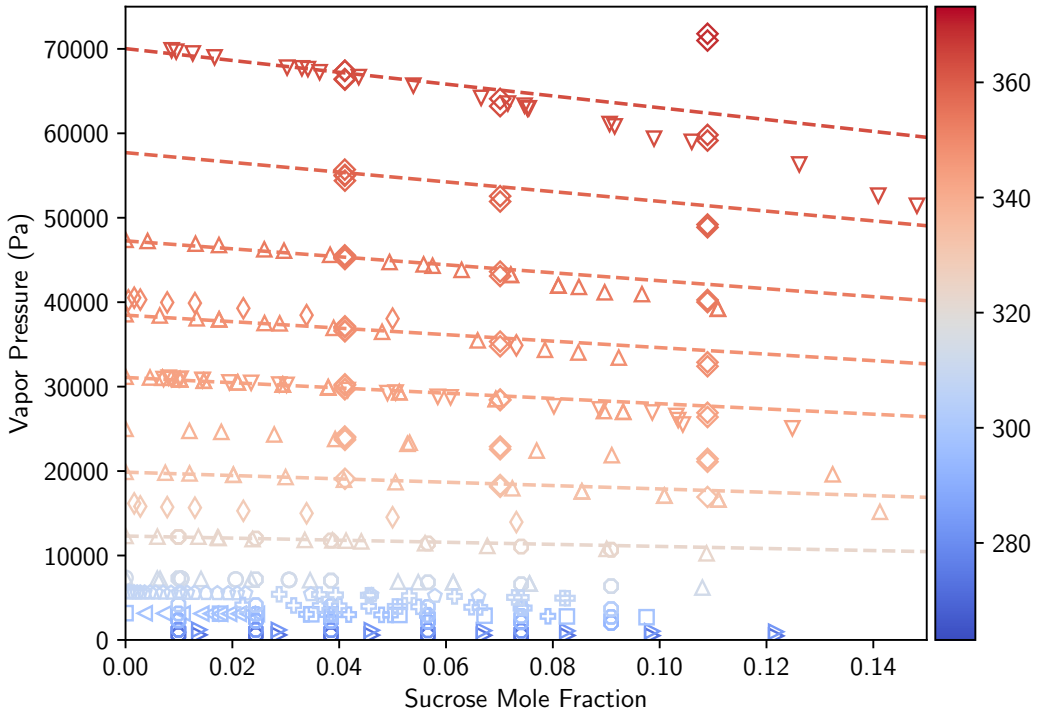


Figure 5.8: Vapor pressure measurements reported in literature for sucrose-water solutions. Open symbols correspond to measurements [168, 242, 164, 165, 172, 243, 244, 150, 245, 246, 167]. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red). Dashed lines indicate ideal (with respect to Raoult's Law) vapor pressure.

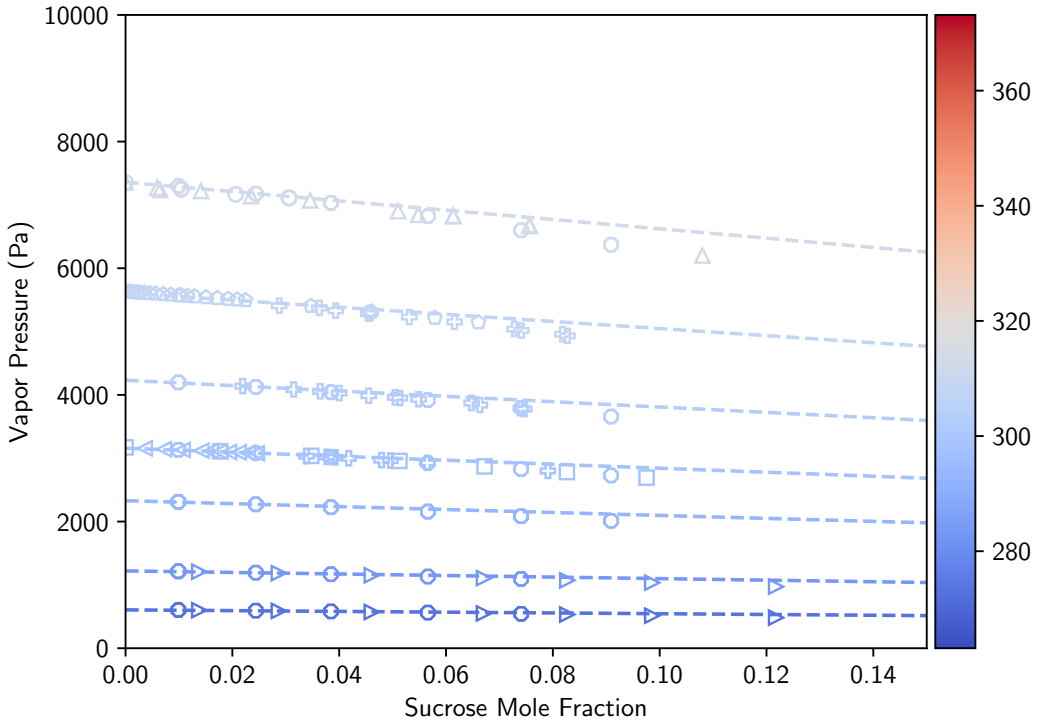


Figure 5.9: Vapor pressure measurements reported in literature for sucrose-water solutions between 273.15 and 323.15 K. Open symbols correspond to measurements [168, 242, 164, 165, 172, 243, 244, 150, 245, 246, 167]. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red). Black dashed line indicates ideal (with respect to Raoult's Law) vapor pressure at 298.15 K.

From Fig. 5.8 and 5.9 it appears all data sets are in good agreement with each other. Further, there appears to be reasonable coverage across the sucrose-water mixture composition and temperature range; however, data at elevated temperatures is limited to five independent reports, while the majority of data points are closer to ambient conditions. With reference to the 'ideal solution' vapor pressure curve (defined relative to Raoult's Law), vapor pressure data indicates the sucrose-water mixture becomes increasingly non-ideal as sucrose composition increases towards saturation (i.e., approximately 0.1 mole fraction), at 298.15 K, inferred from the increasing divergence between dashed lines and symbols in Fig 5.8. The pure vapor pressure (i.e., dashed lines in Fig. 5.8 and 5.9) was estimated using Eq. (5.7), where the vapor pressure of pure water vapor was estimated with the following equation:

$$\log p^{\circ}(T) = 8.07131 - \frac{1730.63}{T + 233.426} \tag{5.19}$$

where p° is saturated vapor pressure (in Torr) and T is temperature (in °C).

The equilibrium relative humidity is another property through which it is possible to investigate the behavior of sucrose-water liquid mixtures. Briefly, the collection contains measurements from thirteen independent sources from as early as 1910 to as recent as 2018, while numerous apparatus have been described. In some cases, the relative humidity is reported as the measured quantity, while in others, water activity is reported as the measured quantity when certain types of modern measurement apparatus are used. For convenience, all measurements where the method of measurement was through analysis of relative humidity were converted to water activity (defined relative to Raoult's Law) through:

$$a_2 = \frac{\mathrm{ERH}}{100}, \tag{5.20}$$

where the equilibrium relative humidity (ERH) was reported as a %, and:

$$a_2 = \frac{h}{100}, \tag{5.21}$$

where the relative pressure (i.e., $h = p_2/p_2^\circ(T)$ with p_2 being the partial pressure of water and $p_2^\circ(T)$ being the saturated vapor pressure of pure water) was reported as %.

All measurements are presented in Fig. 5.10. As shown, it appears all data sets are in qualitative agreement with each other; however, appear somewhat noisy.

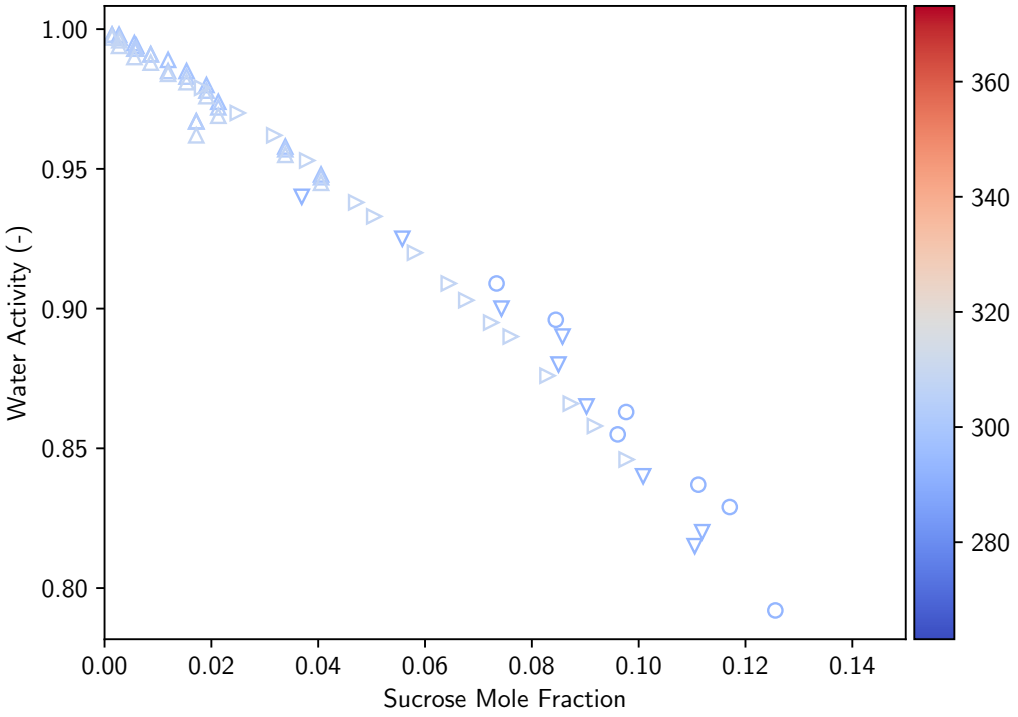


Figure 5.10: Water activity (defined relative to Raoult's Law) in sucrose-water solutions derived from relative humidity-based measurements reported in literature Open symbols correspond to measurements [166, 169, 173, 163]. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red).

The isopiestic molality of sucrose-water solutions relative to different reference solutions (i.e., sodium chloride-water and potassium chloride-water) have been reported by several authors. Briefly, the collection contains measurements from six independent sources from as early as 1938 to as recent as 2021, while numerous apparatus have been described. All measurements are presented in Fig. 5.11 and Fig. 5.12 for sodium chloride and potassium chloride, respectively. As shown, all data sets where measurements are taken relative to the same reference solution are in good agreement with each other. However, it should be noted that all data is reported at 298.15 K.

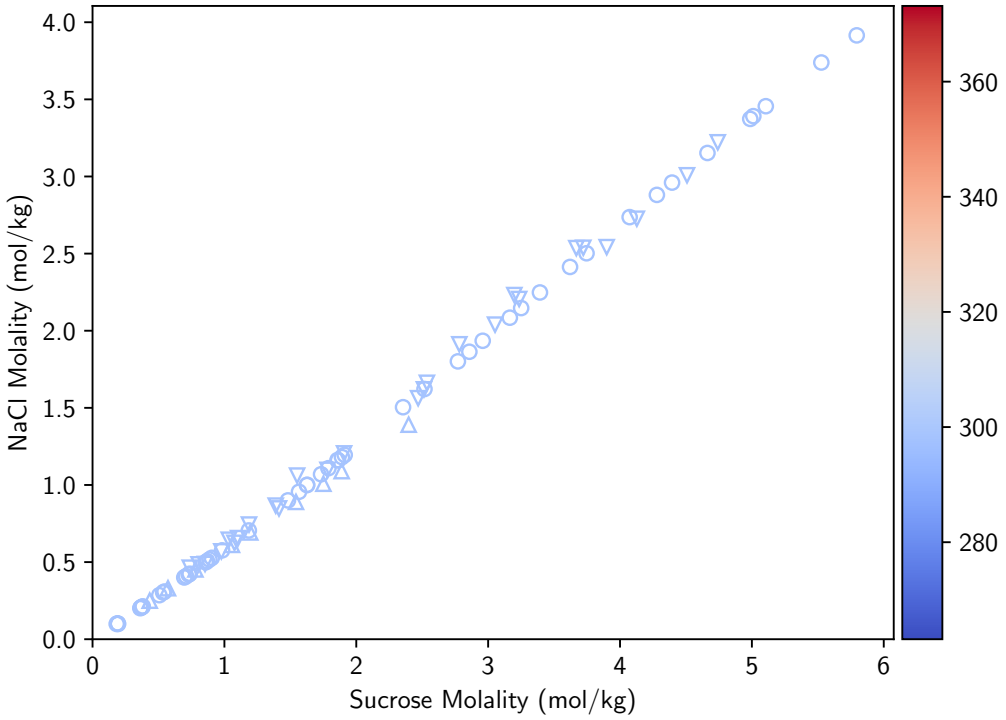


Figure 5.11: Isopiestic molality measurements reported in the literature for sucrose-water liquid mixtures relative to sodium chloride-water liquid mixtures, at 298.15 K. Open symbols correspond to measurements [247, 174, 248, 170].

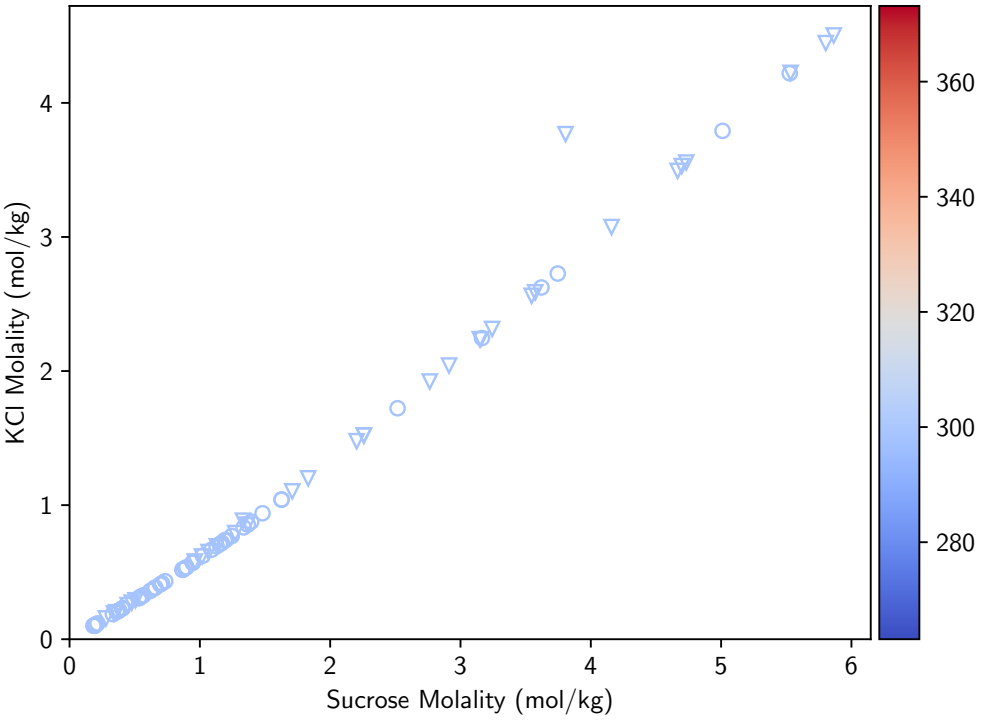


Figure 5.12: Isopiestic molality measurements reported in the literature for sucrose-water liquid mixtures relative to potassium chloride-water liquid mixtures, at 298.15 K. Open symbols correspond to measurements [249, 171, 170].

To convert the isopiestic molality measurments to osmotic coefficients for sucrose-water mixtures, the osmotic coefficient as a function of composition for the relevant reference mixture was used. In this chapter, a reference model for both sodium chloride-water and potassium chloride-water was derived using data tabulated in [250] and [251], respectively, where tabulated data is presented as the output of a large-scale modeling study of both systems based on an extensive collection of experimental thermodynamic data. Both reference models were developed by interpolating available data using the 'RBFInterpolator' function in Scipy with a thin plate spline kernel. The interpolants and data are presented in Fig. 5.13.

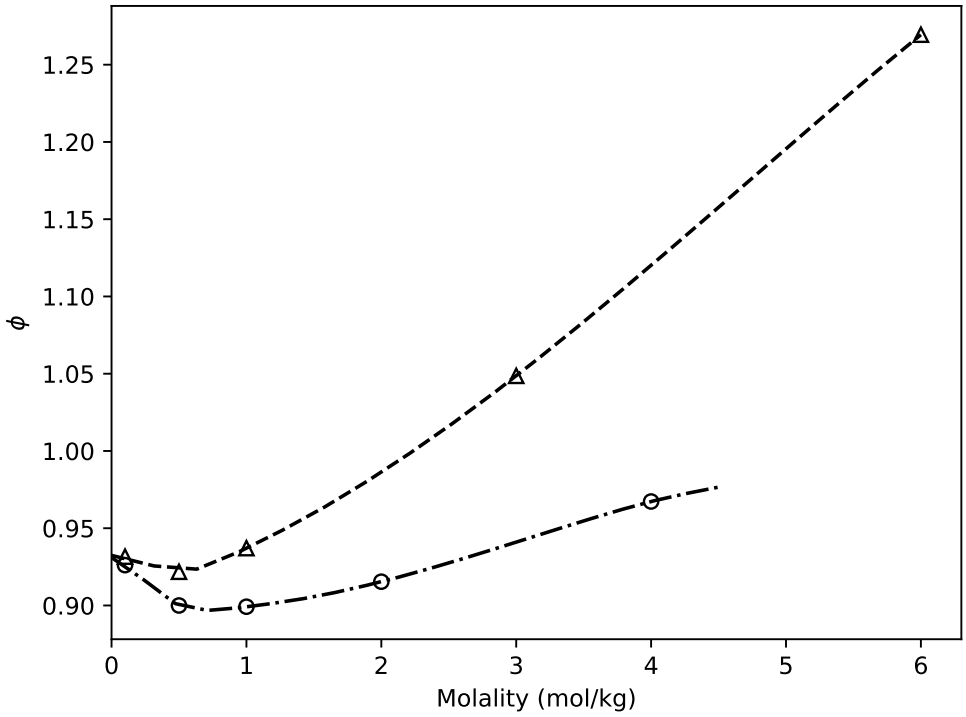


Figure 5.13: Osmotic coefficient as a function of salt molality at 298.15 K. Open symbols correspond to tabualted osotic coefficients for sodium chloride [250] (circles) and potassium chloride [251] (triangles) in water, respectively. Lines correspond to radial basis function interpolants.

The osmotic coefficients derived from isopiestic molality measurements are presented in Fig. 5.14.

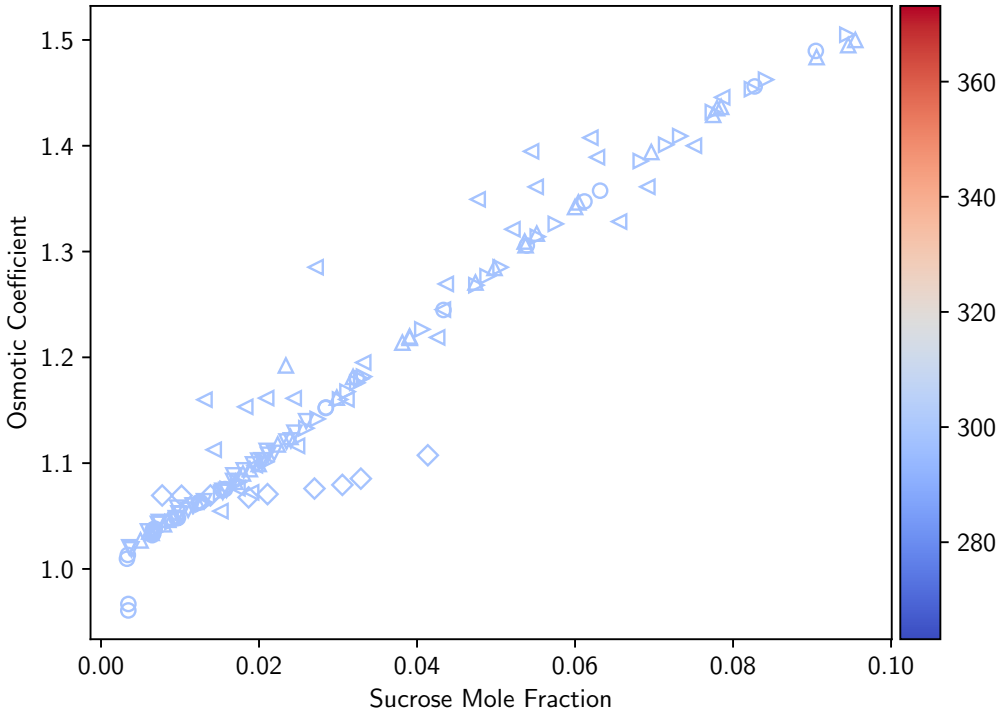


Figure 5.14: Osmotic coefficients for sucrose-water liquid mixtures at 298.15 K. Open symbols correspond coefficients derived from direct measurements reported in literature (NaCl reference [247, 174, 248, 170]. and KCl reference [249, 171, 170].)

As shown, osmotic coefficients estimated from isopiestic molality measurements are in good agreement across the entire range of sucrose composition, with the exception of data reported in [174] (left triangle) and [248] (diamond). It should be noted that two data points in [170] appeared to be outliers when compared to additional points in the data set and are not shown in Fig. 5.14.

All the aforementioned properties give information about the chemical potential of water in the sucrose-water liquid mixture. As such, it is possible to convert each class of measurement into each other, to allow comparison between properties. Here, the measurements derived from relative humidity and isopiestic molality measurements are converted into equivalent vapor pressure. By definition, the osmotic coefficient for a solution comprising a single solute and solvent is related to the solvent activity via:

$$\phi = -\frac{\ln a_2}{\nu M_2 m_1} \tag{5.22}$$

where a_2 is solvent activity, ν is the stoichiometric number of the solute (i.e., ions per molecule) m_1 is solute molality and M_2 is solvent molecular weight (kg/mol). From solvent activity, the equivalent vapor pressure assuming the vapor phase comprises pure solvent is given via the so-

called extended Raoult's Law:

$$p^{\text{vap}}(T) = a_2 p^{\circ,\text{sat}}(T) \tag{5.23}$$

where $p^{\circ,\text{sat}}(T)$ is the pure solvent vapor pressure taken for water as the correlation given in Eq. (5.19).

The combined activity data in terms of water vapor pressure are presented in Fig. 5.15.

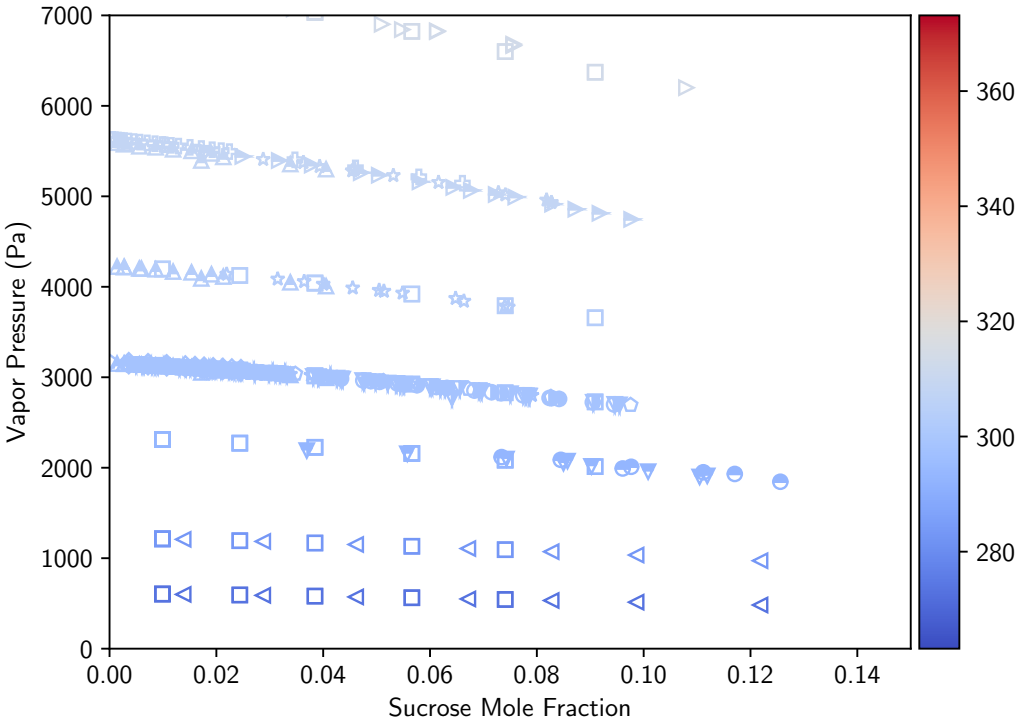


Figure 5.15: Vapor pressure, relative humidity and isopiestic molality measurements reported in literature (processed into equivalent vapor pressure) for sucrose-water solutions between 273.15 and 323.15 K. Open symbols correspond to vapor pressure measurements [168, 242, 164, 165, 172, 243, 244, 150, 245, 246, 167]. Partially closed (top) symbols correspond to vapor pressure estimated from relative humidity measurements [166, 169, 173, 163]. Partially closed (right) symbols correspond to vapor pressure estimated from isopiestic molality measurements [249, 171, 170, 247, 174, 248]. Color gradient corresponds to reported temperature scaled between 260 and 370 K (blue and red, respectively).

It should be noted that the data is shown between 273.15–323.15 K given that this is the range over which measurements of the three sub-classes used in this analysis cover. As shown, measurements from all classes are in good agreement when presented at pure water vapor pressure.

Data derived from relative humidity measurements and isopiestic molalities are compared on the basis of $\ln \gamma_2$ in Fig. 5.16.

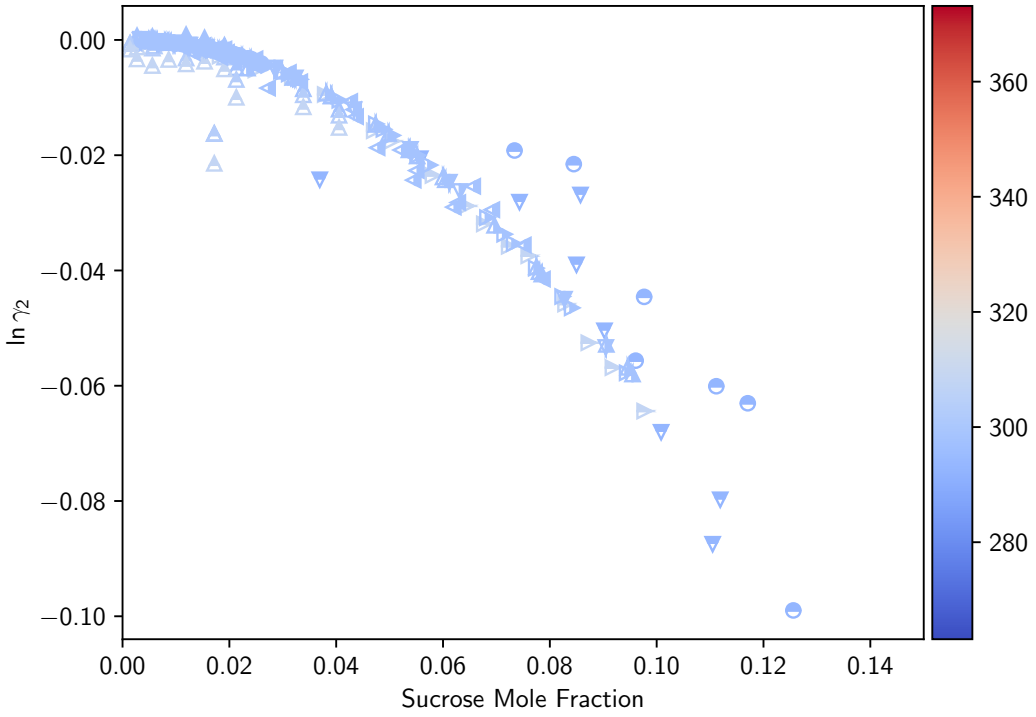


Figure 5.16: Water activity coefficient (i.e., $\ln \gamma_2$), defined relative to Raoult's Law derived from measurements of relative humidity and isopiestic molality measurements reported in literature for sucrose-water solutions between 273.15 and 323.15 K. Partially closed (top) symbols correspond to $\ln \gamma_2$ estimated from relative humidity measurements [166, 169, 173, 163]. Partially closed (right) symbols correspond to $\ln \gamma_2$ estimated from isopiestic molality measurements [249, 171, 170, 247, 174, 248]. Color gradient corresponds to reported temperature scaled between 260 and 370 K (blue and red, respectively).

As shown, estimates of $\ln \gamma_w$ based on relative humidity measurements appear more noisy than those based on isopiestic molalities, with the exception of those reported by [163] (right pointing triangles, top filled), which appear to be in within the same scatter as those based on isopiestic molalities.

5.5.4 Enthalpy of Solution and Dilution

The enthalpy of solution has been reported by various authors. References cited in this work are summarised in Fig. 5.17. As shown, all reports focus on the region of extremely dilute sucrose concentration (i.e., below 0.07 mole fraction, noting that saturation corresponds to 0.1 mole fraction at 298.15 K). In this region and based on available data, the enthalpy of solution appears to have no dependence on composition; however, increases with temperature. It should be noted

that data taken from Barry (1920), Hendricks et al., (1930) as well as Miller and Pablo (2000) was assigned a solution concentration of zero (i.e., infinite dilution) as the true measurement concentration was not found.

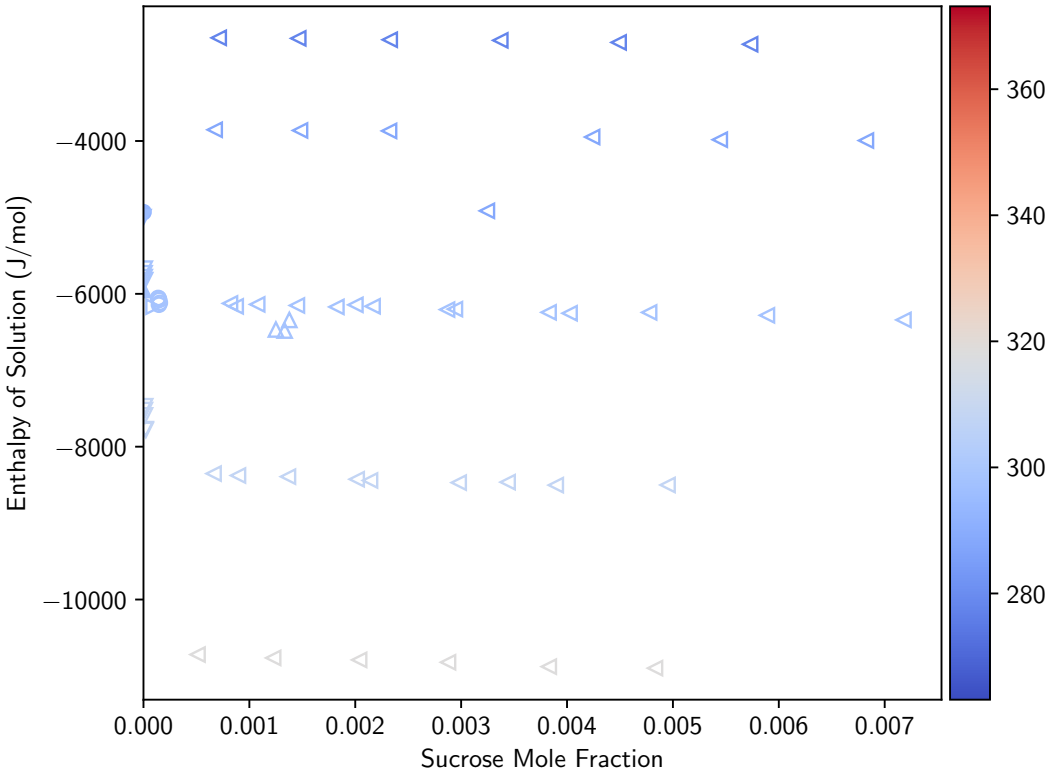


Figure 5.17: Enthalpy of solution data for sucrose-water liquid mixtures reported in literature. Open symbols correspond to measurements [252, 253, 254, 255, 256, 257, 258, 259]. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red).

Data related to the heat evolved during dilution of sucrose-water solutions has been reported by several authors As described in Chapters 3 and 4, the enthalpy of dilution is often reported in various ways. For sucrose-water mixtures at 298.15 K (discussed in Chapter 4), this included two different types of integral measurements reported by [176, 175]. For temperatures other than 298.15 K all additional data found are reported as differential, where a small but finite amount of water is added to a sucrose-water solution of known concentration and the heat evolved measured, giving $\bar{h}_2 - h_2^{l,\circ} = \partial \ln \gamma_2 / \partial \beta$.

Values reported as differential enthalpy of dilution measurements are illustrated in Fig. 5.18. It should be noted that, although reported as integral measurements, differential measurements derived from integral measurements are provided by [176], and also shown. All measurements suggest that the magnitude of differential enthalpy of dilution increases with sucrose mole fraction, as well as temperature and appear to be in good agreement with each other (below approximately

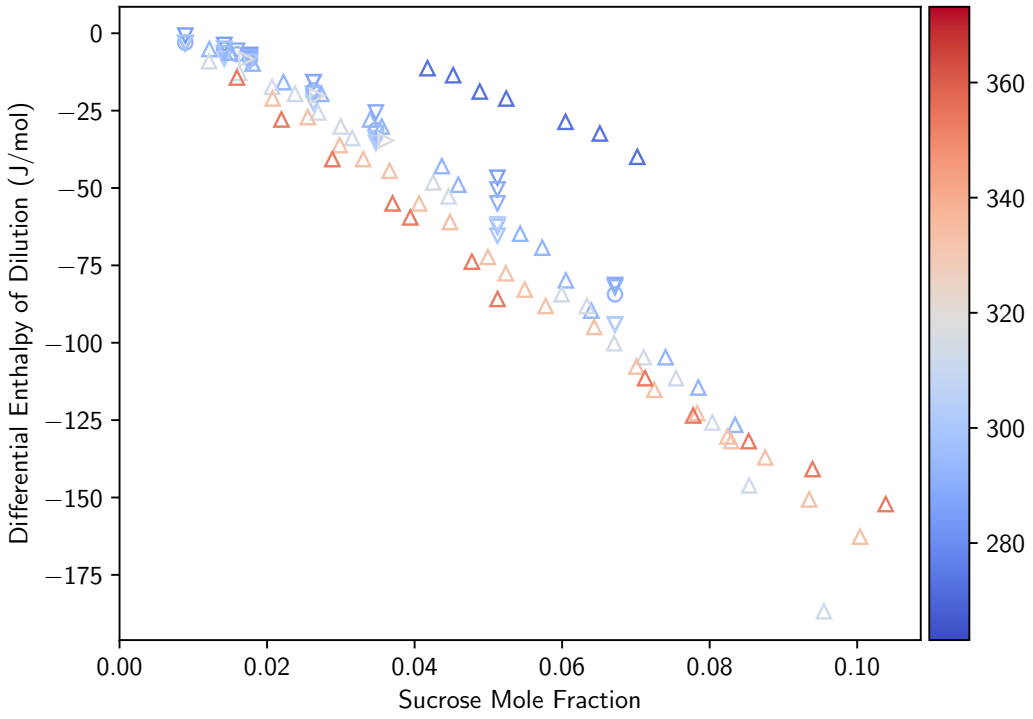


Figure 5.18: Differential enthalpy of dilution data for sucrose-water liquid mixtures reported in literature. Open symbols correspond to measurements [260, 176, 261, 162]. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red).

0.06 mole fraction). However, as also discussed in [145], data reported by [261] at 333.15 and 353.15 K follow the same trend as those at lower temperatures; however, appear to go through an inflection point and begin to level.

5.5.5 Solution Heat Capacity

The heat capacity of sucrose-water liquid mixtures has been reported by various authors, using a variety of techniques. Data are illustrated in Fig. 5.19.

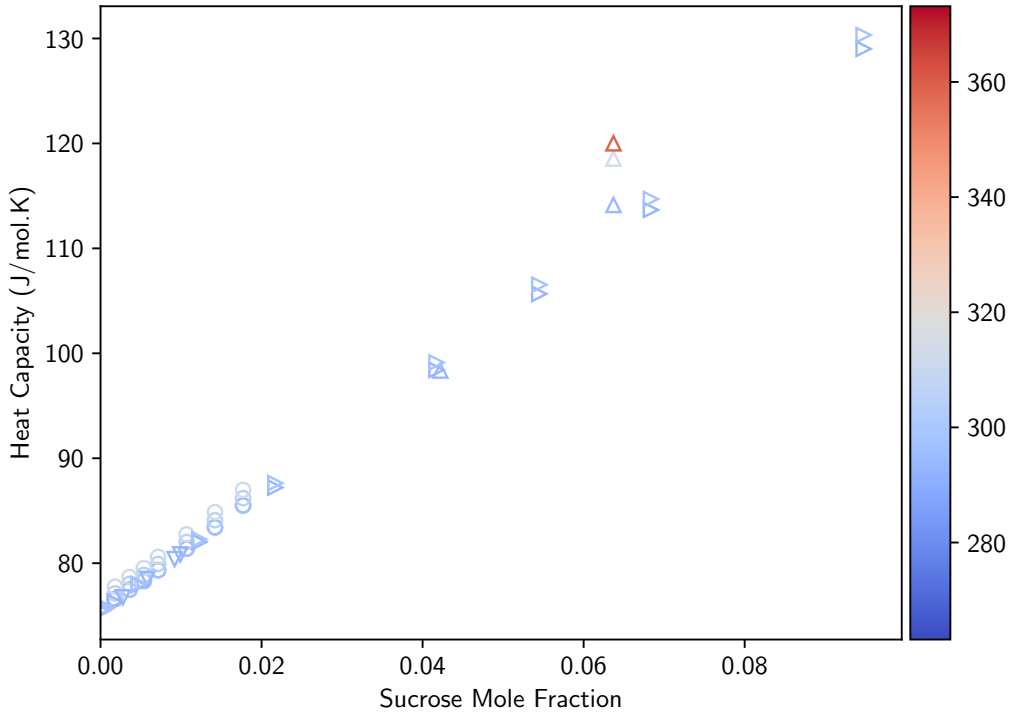


Figure 5.19: Solution heat capacity data for sucrose-water liquid mixtures reported in literature. Open symbols correspond to measurements [179, 177, 178, 180]. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red).

As shown, datasets with coverage of a broad range of sucrose-water liquid mixture compositions is restricted to temperatures at (or near) 298.15 K. There are also two measurements at elevated temperatures (i.e., 330 and 360 K) both at approximately 0.07 sucrose mole fraction.

5.6 Pure Component Reference Values

As noted already, at fixed pressure, the reduced molar Gibbs free energy of single component phase at temperature T can be related to the reduced Gibbs free energy at temperature T_0 via Eq. (5.17). Expressing the Gibbs free energy (and chemical potential) of a pure phase in this way is convenient because, given the heat capacity of all stable components as a function of temperature, as well as the temperature and enthalpy change of all first-order phase transitions in the range of interest, it is possible to determine this model.

For the results reported in this chapter, the Gibbs free energy of crystalline sucrose, as well as liquid water are required to make solubility predictions. relevant pure phases in the sucrose-water system are water (solid, liquid and vapor) and sucrose (solid). As noted already, reference values for the Gibbs free energy of pure phases have been generated and subsequently used in the CALPHAD method by noting that if the enthalpy of formation of elements in their stable states at standard conditions are taken as zero, an interconnected reference state to evaluate 'absolute' Gibbs free energies of compounds is defined. This is shown by first defining the enthalpy of formation of a arbitrary component i is as:

$$\Delta H^{f,\circ} = H_i^\circ - \sum_j b_j h_j, \tag{5.24}$$

where b_j is the stoichiometric amount of element j in component i. The enthalpy of formation is not directly measurable; however, can be found by constructing a hypothetical reaction using Hess's Law, where the enthalpy of combustion is used.

From Postulate IV (see Chapter 2), which states that entropy is zero for a thermodynamic system at zero temperature on the Kelvin scale, it is possible to estimate the entropy of formation. More specifically, this is captured in the Third Law which states that the entropy of a pure crystal is zero at zero temperature, again on the Kelvin scale. At constant pressure, changes in entropy at related to constant pressure heat capacity, giving:

$$S(T) - S(T_0) = \int_{T_0}^{T} \frac{C_p(T)}{T} dT, \tag{5.25}$$

which, for the case where $T_0 = 0\,\mathrm{K}$ becomes:

$$S(T) = \int_{0}^{T} \frac{C_p(T)}{T} dT, \tag{5.26}$$

which makes it possible to estimate entropy with heat capacity low temperature heat capacity. The entropy of formation is defined in the same way as enthalpy:

$$\Delta S^{f,\circ} = S_i^\circ - \sum_j b_j S_j^\circ. \tag{5.27}$$

Using Eq. (5.24) and (5.27), the standard Gibbs free energy of formation is found from:

$$\Delta G^{f,\circ} = \Delta H^{f,\circ} - T\Delta S^{f,\circ}. \tag{5.28}$$

The relevant properties are tabulated for many compounds in property handbooks [262]. This makes it possible to calculate the standard Gibbs free energy of formation for a pure component in it's stable state, giving a consisent reference on which calculations are based.

This, coupled with calorimetric data (i.e., pure component heat capacity as a function of temperatre in solid, liquid and gas states, as well as temperature and enthalpy of first-order phase transitions, all of which are experimentally measurable in-principle) the Gibbs free energy at any temperature and fixed pressure can be calculated directly from fundamental thermodynamic relations.

Standard properties of pure crystalline sucrose and pure liquid water reported in literature are presented in Table 5.3. The standard Gibbs free energy for pure sucrose is not available in

Table 5.3: Gibbs free energy and enthalpy for pure crystalline sucrose and liquid water at 1 atm and 298.15 K defined relative to the elements in their standard state at same temperature and pressure. † denotes properties estimated from standard values reported by NIST

Compound	g° (kJ/mol)	h° (kJ/mol)
Sucrose$^{(s)}$	-1553.16	-2221.2†
Water$^{(l)}$	-237.14†	-285.8†

the literature; however, the value reported in Table 5.3 was evaluated using Eq. (5.28), where $\Delta H^{f,\circ}$ was taken as the value in Table 5.3, T was taken as 298.15 K and $\Delta S^{f,\circ}$ was estimated using 392.4, 130.68, 205.15 and 5.6 J/mol for the standard entropy of sucrose (crystal), hydrogen (gas), oxygen (gas) and carbon (solid), respectively (where $\Delta S^{f,\circ}_{\text{sucrose}} = S_{\text{sucrose,298.15 K}} - (12 S_{\text{carbon,298.15 K}} + 11 S_{\text{hydrogen,298.15 K}} + 5.5 S_{\text{oxygen,298.15 K}})$ and all values are reported by NIST).

5.7 Multi-component Phase Modelling and Phase Equilibria Prediction

In the previous section, Gibbs free energy models were generated for thermodynamically stable phases observed in the sucrose-water system (i.e., pure crystalline sucrose, as well as pure solid, liquid and vapor water). In this section, models for the Gibbs free energy of sucrose-water liquid mixtures are constructed, and the solubility of sucrose in water is subsequently predicted. The section starts with a detailed discussion on available experimental thermodynamic data presented previously with the goal of highlighting the temperature and composition dependence of the sucrose-water liquid solution. Given that constructing a Gibbs free energy model for the liquid mixture is more complex than for single components where only a single type of experimental data

is required (i.e., heat capacity), a short discussion on data modelling is provided. Then, modelling results are presented for the 'full' experimental database described in the previous section. Modelling is completed based on different models, all of which are derived from the Redlich-Kister polynomial (for the excess contribution), and a contribution in the ideal Gibbs free energy associated to the hypothetical pure liquid sucrose that yields a heat capacity with a linear temperature dependence.

5.7.1 Discussion on Experimental Data

In this section, the construction of a Gibbs free energy model for the sucrose-water liquid mixture is disscussed in the specific context of modelling the experimental data presented in Section 5.5. As shown, several different thermodynamic quantities have been reported in the literature, including those related to the solvent activity (e.g., vapor pressure, isopiestic molality, relative humidity and solution boiling point), changes in enthalpy (e.g., dilution and solution), as well as solution heat capacity. Assuming all measurements are correct, they should be thermodynamically consistent. Before developing a model for the Gibbs free energy of sucrose-water liquid mixtures, a discussion on the experimental data is provided, with a view to identify key characteristics any set of Gibbs free energy models should have to adequately describe it.

As described already, interpreting the activity (or, more specifically, activity coefficient) of solvent in multi-component liquid mixtures (typically through vapor-liquid equilibrium) is one of the primary methods used to construct thermodynamic models. Based on the experimental data illustrated in Fig. 5.16, there are several independent reports of solvent activity at 298.15 K, the majority of which are in good agreement (likely driven by the use of isopiestic molality measurements), and indicate the solvent activity coefficient (defined with respect to Raoult's Law) has a quadratic dependence. Furthermore, although noisy, solvent activity coefficient data at temperatures other that 298.15 K indicates some form of temperature dependence which is in qualitative agreement with enthalpy of dilution data discussed next. However, it is challenging to discern the nature of this due to the lack of more accurate isopiestic molality measurements at temperatures other than 298.15 K, with all other data derived from seemingly less accurate relative humidity and vapor pressure measurements.

Focusing on the differential enthalpy of dilution data shown in Fig. 5.18, there are four independent reports at various temperatures and composition ranges. As mentioned already, composition dependent data reported in [260], [176] and [261] appear to be in good agreement at temperatures near 298.15 K. More specifically, the data appear to follow a consistent dependence on composition, irrespective of temperatures, as well as having a dependence on temperature. As an aside, this is potentially in contrast to modelling described in Chapter 4, where the solution heat capacity was assumed to be a linear function of composition (i.e., no excess contribution).

However, data reported in [261] includes additional measurements at 273.15, 333.15 and 353.15 K where there are no other data at these temperatures from independent sources. In terms of constructing a model for the Gibbs free energy of sucrose-water liquid mixtures, this is important because data reported in [261] at 333.15 and 353.15 K appear to have a different composition dependence to data at temperatures around 298.15 K discussed already, as illustrated in Fig. 5.18. This behavior was discussed in [145], where the authors suggest that 'differential enthalpy of dilution measurements are less accurate than integral measurements', while noting that 'calculated mean differential heats of dilution reported in [261], involving rare measurements of differential heats of dilution, was at least 10–15% based on available integral measurement data' and reiterating that the authors in [261] describe challenges with temperature control due to viscous nature of sucrose-water liquid mixtures at high temperature and composition. As an aside, it should be noted that it is unclear how the authors estimated errors in the reported differential enthalpy of dilution measurements. In the original work [261], the authors also discuss the temperature and composition dependence of the reported differential enthalpy of dilution measurements, commenting that a similar behavior was observed for differential enthalpy of dilution of hydrochloric acid in water [263].

As noted already, reports of thermodynamic properties should be consistent if they are accurate. In theory, differential enthalpy of dilution measurements are related to solvent activity measurements, since:

$$\bar{h}_2^{\mathrm{l}}(T,x) - h_2^{\mathrm{l},\circ}(T) = \frac{\partial \ln \gamma_2(T,x)}{\partial \beta} \tag{5.29}$$

where all symbols retain their definitions discussed already in this chapter. As noted in Section 5.5, there are four reports of water vapor pressure measurements above sucrose-water liquid mixtures at temperatures between 313.15–363.15 K and 0–0.2 sucrose mole fraction [168, 242, 164, 150]. To provide rationale for either including or rejecting high temperature differential enthalpy of dilution measurements reported in [261], it was thought that this data could be used to estimate the differential enthalpy of dilution. However, as shown in Fig. 5.20, when converted from vapor pressure to water activity coefficient, the data is subject to significant noise and discerning a meaningful relationship is challenging. It should be noted that a detailed description of calculations can be found in Appendix D.

Focusing on the enthalpy of solution data illustrated in Fig. 5.17, this data indicates that the magnitude of sucrose enthalpy of solution in water increases with increasing temperature (e.g., approximately 4,000 J/mol at 278.15 K and approximately 12,000 J/mol at 318.15 K, for data reported in [259]). Based on this data set (i.e., [259]), it appears that the rate of change of the enthalpy of solution (at concentrations assessed) varies in temperature as noted by increasing differences in the reported enthalpy of solution for similar increases in temperature; however, it should be noted that there are limited reports available to support these measurements. Given

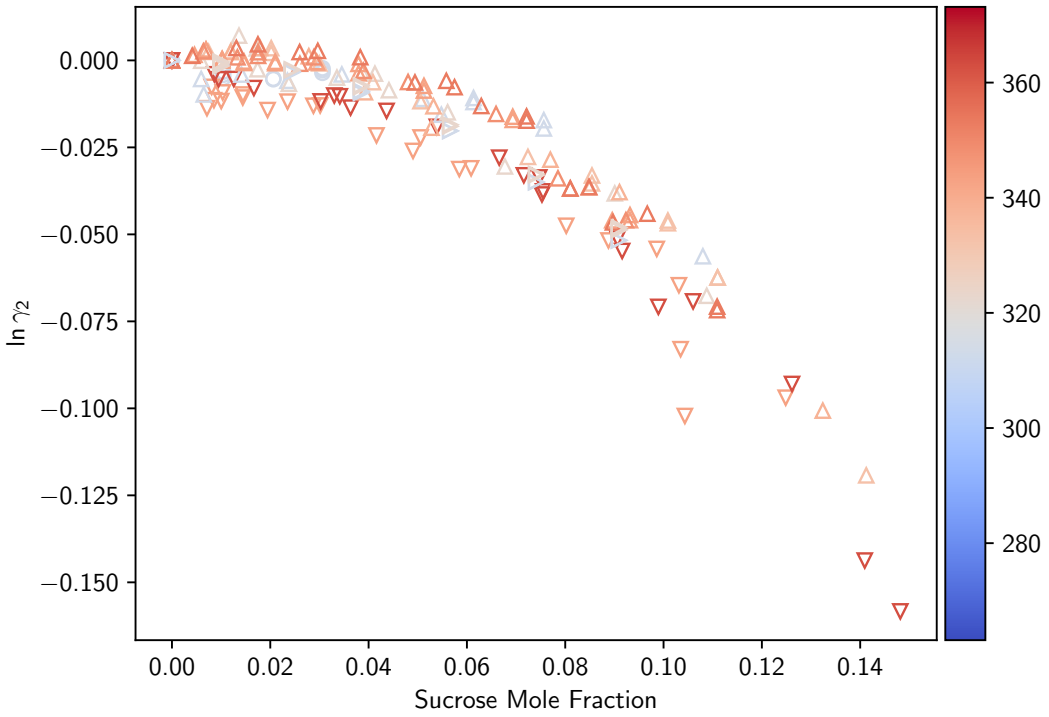


Figure 5.20: Water activity coefficients in sucrose-water liquid mixtures estimated from vapor pressure measurements reported between 313.15–373.15 K. Open symbols correspond to experimental measurements. Color gradient corresponds to reported temperature scaled between 260 K (blue) and 370 K (red).

that all enthalpy of solution data collated in this chapter is restricted to highly dilute sucrose-water liquid mixtures, with 0.007 sucrose mole fraction being the most concentrated data point, data at temperatures other than 298.15 K was omitted from fitting. To re-cap, from a theoretical perspective the enthalpy of solution is defined as:

$$\begin{aligned}\Delta H^{\text{sol.}} &= n_1\left(h_1^{\text{l},\circ} + \bar{h}_1^{\text{E}}\right) + n_2\left(h_2^{\text{l},\circ} + \bar{h}_2^{\text{E}}\right) - n_1 h_1^{\text{s},\circ} - n_2 h_2^{\text{l},\circ} \\ &= n_1\left(\Delta h^{\infty,s} - h_1^{\text{l},\infty}\right) + n_1\bar{h}^{\text{E}} + n_2\bar{h}_2^{\text{E}}\end{aligned} \tag{5.30}$$

which reduces to:

$$\Delta H^{\text{sol.},\infty} = n_1\left(\Delta h_1^{\infty,s}(T) - h_1^{\text{l},\infty}(T)\right) \tag{5.31}$$

for infinitely dilute solutions (noting that all symbols are defined in the same way as Chapter 4), which must have the required functional form to capture the experimentally observed behavior.

Although limited, available solution heat capacity experimental data (illustrated in Fig. 5.19) appears to have a temperature dependence where, at fixed solution composition, the solution heat capacity increases with increasing temperature In terms of constructing a Gibbs free energy model, there are different model contributions that would allow this behaviour including the term associated to the hypothetical pure liquid solute, the excess contribution, or a mixture of both. Assuming the temperature dependence noted in the discussion on differential enthalpy of dilution data is real, this implies that the chosen Gibbs free energy model used to describe the sucrose-water liquid mixture should have an excess contribution component for the heat capacity. Interestingly, this is in contrast to the modelling approach used in Chapter 4, where it was assumed that the solution heat capacity could be modelled with 'ideal' terms, only.

5.7.2 Data Modelling

As described already, constructing Gibbs free energy models typically involves finding model structures and associated parameter values that are in 'best' agreement with available experimental data. However, it is often the case that the numerical values of different types of experimental data used during model development are drastically different partially due to the reported units of measurement. For example, vapor pressure measurements for sucrose-water liquid mixtures span 1000–70,000 Pa, while osmotic coefficients derived from isopiestic molality measurements span 1–1.1.

Furthermore, it is common practice to frame Gibbs free energy model parameterisation as a single-criteria optimisation where there is a single objective function comprising error contributions from each data type. However, if the contribution to the objective function for each data type was interpreted as the sum of errors squared, it is straightforward to see that different data types will contribute different errors to the objective function driven predominantly by the magnitude of the numerical values. It is unclear how ignoring this mismatch in contributions to the objective

function will impact numerical optimisation of parameters, as such, it is common to scale data based on different features such as estmated uncertainty and/or expert judgment [198].

When the goodness-of-fit measure is taken as the sum of errors squared, this gives the an objective function of the form:

$$F = \sum_{i}^{N} (\hat{y}_i(\theta, x_i) - y_i(x_i))^2 \cdot p_i \tag{5.32}$$

where $\hat{y}_i(\theta, x_i)$ is thermodynamic quantity y predicted using the current parameter estimate θ at conditions x_i, y_i is thermodynamic quantity y measured at conditions x_i, N is the number of data points, and p_i is a 'weighting factor' [198], which in practice may take different forms. For example, in some cases $p_i = 1/\sigma_i^2$ where σ_i corresponds to an estimate of data uncertainty for measurement i. Alternatively, $p_i = w_i/\sigma_i^2$ where w_i is an additional 'weighting factor' used to augment the contribution of data during optimisation. It should be noted that Eq. (5.32) is readily derived by assuming data is normally distributed and is a special case the so-called 'generalised least squares' theory; however, this contribution is usually omitted when data of the same type is used under the assumption that any associated uncertainty is independent and identically distributed (i.i.d). It should be noted that this type of objective function is applied in studies outside the scope of CALPHAD assessments [145, 80].

As an aside, and noted briefly already, the correlation of different types of experimental thermodynamic data for the construction of mathematical models describing the Gibbs free energy of phases is usually framed as a single criteria optimisation problem, However, recent work described in [213] has shown that interpreting the problem as a multi-criteria optimisation (MCO) can yield robust results as illustrated for the parameterisation of an NRTL-based excess Gibbs free energy model describing binary liquid mixtures containing water, 1-propanol and 2-pentanol.

For the purpose if this chapter, given the lack of robust information on possible uncertainty, each data type was assigned a specific value for σ (i.e., all vapor pressure measurements have the same assumed data uncertainty, irrespective of the technique used or reporting authors), and those values are tabulated in Table 5.4.

Table 5.4: Estimated uncertainty for sucrose-water thermodynamic properties.

Quantity	σ (units)
Vapor Pressure	100 (Pa)
Osmotic Coefficient	0.005 (-)
Differential Enthalpy of Dilution	15 J/mol
Integral Enthalpy of Dilution	15 J/mol
Solution Heat Capacity	1 J/mol.K

As shown in Table 5.4 vapor pressure measurements were assumed to have a relative uncertainty of 100 Pa, which was estimated from data reported in [164], [243] and [150], where each report includes at least two measurements reported at identical (or similar very similar) conditions which, generally speaking, were within 100 Pa, with the exception of some points in [243] which differed by almost 1000 Pa. Furthermore, analysis of data from independent sources reporting measurements at 298.15 K indicated that measurements at similar sucrose compositions were generally within 100 Pa. Osmotic coefficients derived from isopiestic molality measurements were assigned an uncertainty of 0.005, based on the variation between independent reports at 298.15 K across a broad range of sucrose compositions. Integral and differential enthalpy of dilution measurements were both assigned an arbitrary uncertainty of 15 J/mol. Solution heat capacity measurements were assigned an arbitrary uncertainty of 1 J/mol.K. In addition, all data sets were weighted the same.

5.7.3 Redlich-Kister Modelling

As noted already, the Redlich-Kister polynomial is used to model the excess Gibbs free energy contribution in multi-component phases throughout the thermodynamic modelling literature, and extensively in CALPHAD-based modelling studies. Interestingly, based on the review of sucrose-water modelling studies provided earlier in this chapter, no studies have been found where the Redlich-Kister polynomial has been used to model sucrose-water liquid mixtures.

Given the nature of Redlich-Kister polynomial excess Gibbs free energy models, it is not possible to determine an appropriate model structure (and associated parameter values) *a priori*, and this has to be found through trial-and-error by correlating available experimental data. As noted already, parameters in the Redlich-Kister polynomial are assumed to have a temperature dependence modeled by Eq. (5.15); however, as noted in [205], it is rare to use all parameters in an assessment and often only the first three terms are used. In the remainder of this section, modelling the excess contribution in the Gibbs free energy of the sucrose-water mixture is restricted to the first three temperature dependent parameters.

Full Sucrose-water Liquid Mixture Data Set

Given uncertainty around the reliability and thermodynamic consistency of available experimental data, it was initially assumed that all data was accurate and included in the model development process. From this, the parameterisation of Gibbs free energy models for the sucrose-water liquid mixture was approached in a two-step process. In the first step, parameters in the excess Gibbs free energy contribution (described by a Redlich-Kister polynomial), as well as parameters associated to the hypothetical pure liquid sucrose heat capacity were correlated to experimentally measured

quantities reported for the sucrose-water liquid mixture, including solvent activity (i.e., vapor pressure and isopiestic molality), enthalpy of dilution (i.e., differential and integral), and solution heat capacity—all data used in the fitting is captured in Table 5.6.

It should be noted that certain data reported earlier in this chapter was omitted from this study, including: solution heat capacity data reported in [179], [177], and [180] (at concentrations below 0.012 sucrose mole fraction), due to a high number of data points covering a narrow composition range at the same temperature, and potentially biasing fits. In addition, to simplify the fitting procedure, the heat capacity of pure of liquid water water was also treated as a constant, with a value of 75.3 J/mol.K and enthalpy of solution to infinite dilution at 298.15 K was fixed at 6,000 J/mol. Finally, only solution properties were used—boiling and freezing point data were omitted.

In the second step, parameters associated to the reference Gibbs free energy and enthalpy of the hypothetical pure liquid sucrose were determined using the parameters derived in the first step. More specifically, as noted in Section 5.2.2, only temperature dependent solubility data can provide information about the parameter associated to the reference Gibbs free energy of the hypothetical pure liquid sucrose at T_0 (i.e., $g_1^{l,\circ}(T_0)$). In keeping with the general theme of this thesis, this parameter was evaluated using the estimated sucrose-water solubility at 298.15 K, noting that T_0 =298.15 K, and that:

$$g_1^{l,\circ}(T_0) = g_1^{s,\circ}(T_0) + \ln x_{1,\text{sat}}(T_0)\gamma_1(x_{1,\text{sat}}(T_0)), \tag{5.33}$$

where $x_{1,\text{sat}}(T_0)$ is the solubility sucrose in water at T_0 (taken as 0.098 at 298.15 K), $g_1^{s,\circ}(T_0)$ is the reference molar Gibbs free energy of pure crystalline sucrose at T_0 (evaluated in Section 5.6), and γ_1 is the sucrose activity coefficient in the sucrose-water liquid mixture, evaluated from the excess Gibbs free energy model developed in step one. Again as noted in Section 5.2.2, it is possible to evaluate the parameter associated to the reference enthalpy of the hypothetical pure liquid sucrose at T_0 (i.e., $h_1^{l,\circ}(T_0)$) with temperature dependent solubility (at temperatures other than 298.15 K) and/or temperature dependent enthalpy of solution data, only. Based on the availability of several enthalpy of solution measurements at 298.15 K and at infinite dilution, this parameter was found using this value, via:

$$h_1^{l,\circ}(T_0) = h_1^{s,\circ}(T_0) + \frac{\partial \ln \gamma_1^{\infty}(T_0)}{\partial \beta}, \tag{5.34}$$

where $h_1^{s,\circ}(T_0)$ is the molar enthalpy of pure crystalline sucrose at T_0 (evaluated in Section 5.6), and γ_1^{∞} is the estimated infinitely dilution activity coefficient for sucrose in sucrose-water liquid mixtures evaluated from the excess Gibbs free energy model developed in step one.

Three forms of the Redlich-Kister model were correlated against the experimental database, starting with model I, defined as:

$$g^{E} = x_1 x_2 \left[{}^{(0)}a_{ij} + {}^{(0)}b_{ij} T + {}^{(0)}c_{ij} T \ln T \right], \tag{5.35}$$

model II, specified as:

$$g^{\mathrm{E}} = x_1 x_2 \left[{}^{(0)}a_{ij} + {}^{(0)}b_{ij} T + {}^{(0)}c_{ij} T \ln T \right. \\ \left. (x_1 - x_2) \left({}^{(1)}a_{ij} + {}^{(1)}b_{ij} T + {}^{(1)}c_{ij} T \ln T \right) \right], \tag{5.36}$$

and model III, defined as:

$$g^{\mathrm{E}} = x_1 x_2 \left[{}^{(0)}a_{ij} + {}^{(0)}b_{ij} T + {}^{(0)}c_{ij} T \ln T \right. \\ (x_1 - x_2) \left({}^{(1)}a_{ij} + {}^{(1)}b_{ij} T + {}^{(1)}c_{ij} T \ln T \right) \\ \left. (x_1 - x_2)^2 \left({}^{(2)}a_{ij} + {}^{(2)}b_{ij} T + {}^{(2)}c_{ij} T \ln T \right) \right], \tag{5.37}$$

where ${}^{(k)}l_{ij}$ denotes a flexible parameter (with l referring to a, b or c, respectively), as noted previously. For each model, parameter values were derived by finding the minimum of Eq. (5.32), with p_i defined as $1/\sigma_i^2$ (where σ_i for each experimental data type was taken as the associated value in Table 5.4). For this analysis, the Nelder-Mead numerical optimisation algorithm [160] implemented in SciPy [161] was used alongside the Python modules described in Section 5.3. Best fit parameters for based on models I, II, and III to the full experimental data set alongside the 'best-fit' value of Eq. (5.32) are summarised in Table 5.5. As an aside, given that the Nelder-Mead numerical optimisation algorithm implemented in Scipy requires a set of initial values, those used for each parameter are also provided in Table 5.5. The parameter associated to the constant pressure heat capacity of hypothetical liquid sucrose at T_0 (i.e., $c_{p,1}^{l,\circ}(T_0)$) was assumed to be 650 J/mol.K based on work described in Chapter 4 and the associated local temperature dependence (i.e., $c_{p,1}^{\prime l,\circ}(T_0)$) was assumed to be 1.66 J/mol.K^2 which is the same as the constant pressure heat capacity temperature dependence for pure crystalline sucrose (at 298.15 K), as described in Section 5.6. All parameters in the Redlich-Kister excess Gibbs free energy contribution were assumed to be 0.

Table 5.5: Parameter values.

Parameter	Initial	Model I	Model II	Model III	Model I (Reduced)
$c_{p,1}^{l,\circ}(T_0)$ (J/mol.K)	650	721.57	674.41	692.13	731.45
$c_{p,1}^{\prime l,\circ}(T_0)$ (J/mol.K^2)	1.66	2.26	2.25	2.26	2.26
A_0 (J/mol)	0	-71.66	-20.85	6.93	-81.86
A_1 (J/mol)	0	80.71	-11.88	-4.01	19.42
A_2 (J/mol.K)	0	8.27	-0.75	5.12	9.55
B_0 (J/mol)	0	-	-18.76	-10.61	-
B_1 (J/mol)	0	-	16.71	4.53	-
B_2 (J/mol.K)	0	-	3.39	-3.65	-
C_0 (J/mol)	0	-	-	-12.37	-
C_1 (J/mol)	0	-	-	1.07	-
C_2 (J/mol.K)	0	-	-	3.33	-
$g_1^{\circ,l}(T_0)$ (J/mol)	-	-1533304.13	-1503018	-1603218	-1532859
$h_1^{\circ,l}(T_0)$ (J/mol)	-	-2194781	-2208661	-2203303.8	-2191547
Eq. (5.32)	-	1899	996.5	853.15	1632.02

As shown, the magnitude of the objective function (i.e., Eq. (5.32)) decreases as the number of parameters in the Redlich-Kister polynomial increases. Generally speaking, this is expected given that an increased number of parameters gives greater model flexibility and data correlation capacity. Reviewing this metric in isolation implies that model III gives the 'best' fit to available experimental data, with model I being the worst.

To gain additional insight into the fitting process, the 'best fit' objective function (i.e., Eq. (5.32)) is broken down into the contributions per each dataset used in fitting in Table 5.6.

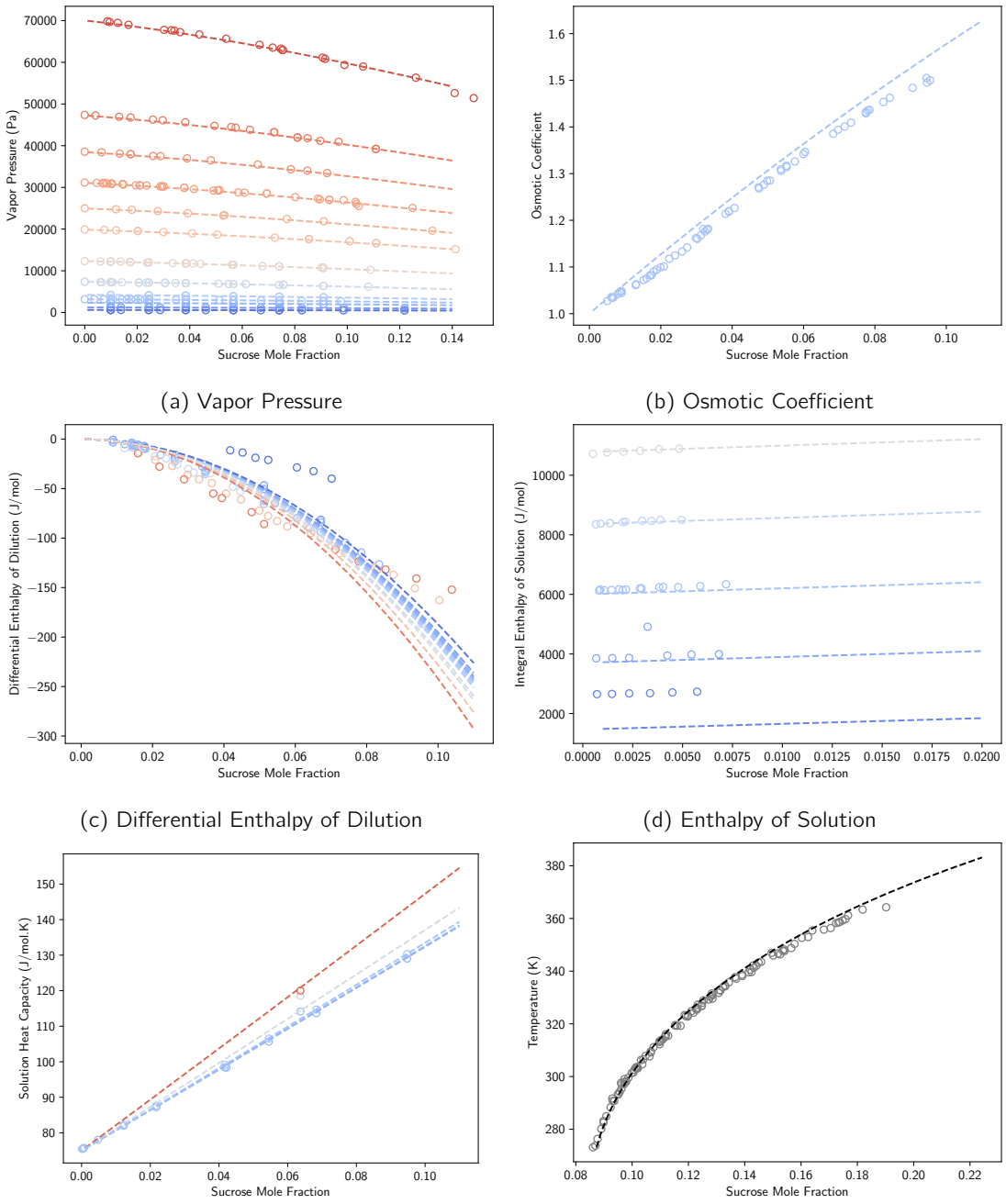
Table 5.6: Objective function contributions broken down by data set

Type	Reference	Model I	Model II	Model III
Vapor Pressure	[168]	0.45	0.44	0.48
	[242]	775.25	293.24	288.58
	[164]	612.12	294.29	269.72
	[165]	0.07	0.05	0.04
	[172]	0.11	0.07	0.05
	[150]	6.41	1.09	1.09
	[245]	0.23	0.06	0.06
Isopiestic Molality	[170]	9.14	27.06	38.05
	[247]	8.87	22.82	33
Differential Enthalpy of Dilution	[260]	0.527	0.54	0.36
	[176]	4.29	3.41	2.61
	[261]	202.37	188.93	130.72
	[162]	0.28	0.12	0.029
Integral Enthalpy of Dilution	[175]	253.45	138.87	56.83
Solution Heat Capacity	[178]	22.97	22.96	22.93
	[180]	2.56	2.56	2.56
Total		1899.1	996.5	853.2

As shown, the reduction in best fit objective function between model I and model II and III, is primarily driven by a substantial reduction in the error associated to two vapor pressure data sets [242, 164], and the integral enthalpy of dilution data set [175]. However, it should be noted that the best fit parameters in model I gave a model in better agreement with the solvent activity data derived from isopiestic molality measurements—perhaps the most reliable of all the data included in the analysis.

Thermodynamic properties derived from best-fit results for models I, II, and III are compared to corresponding experimental data illustrated in Fig. 5.21, 5.22, and 5.23, for models I, II, and III, respectively (sub-figures (a)–(e)). As shown, the best fits to vapor pressure and solution heat capacity gives a model in good agreement with available data across all temperatures and compositions used, for all models.

From Fig. 5.21, model I appears to predict the osmotic coefficient data derived from isopiestic molality measurements in good agreement with experimental data. In contrast, from Fig. 5.22



(a) Vapor Pressure

(b) Osmotic Coefficient

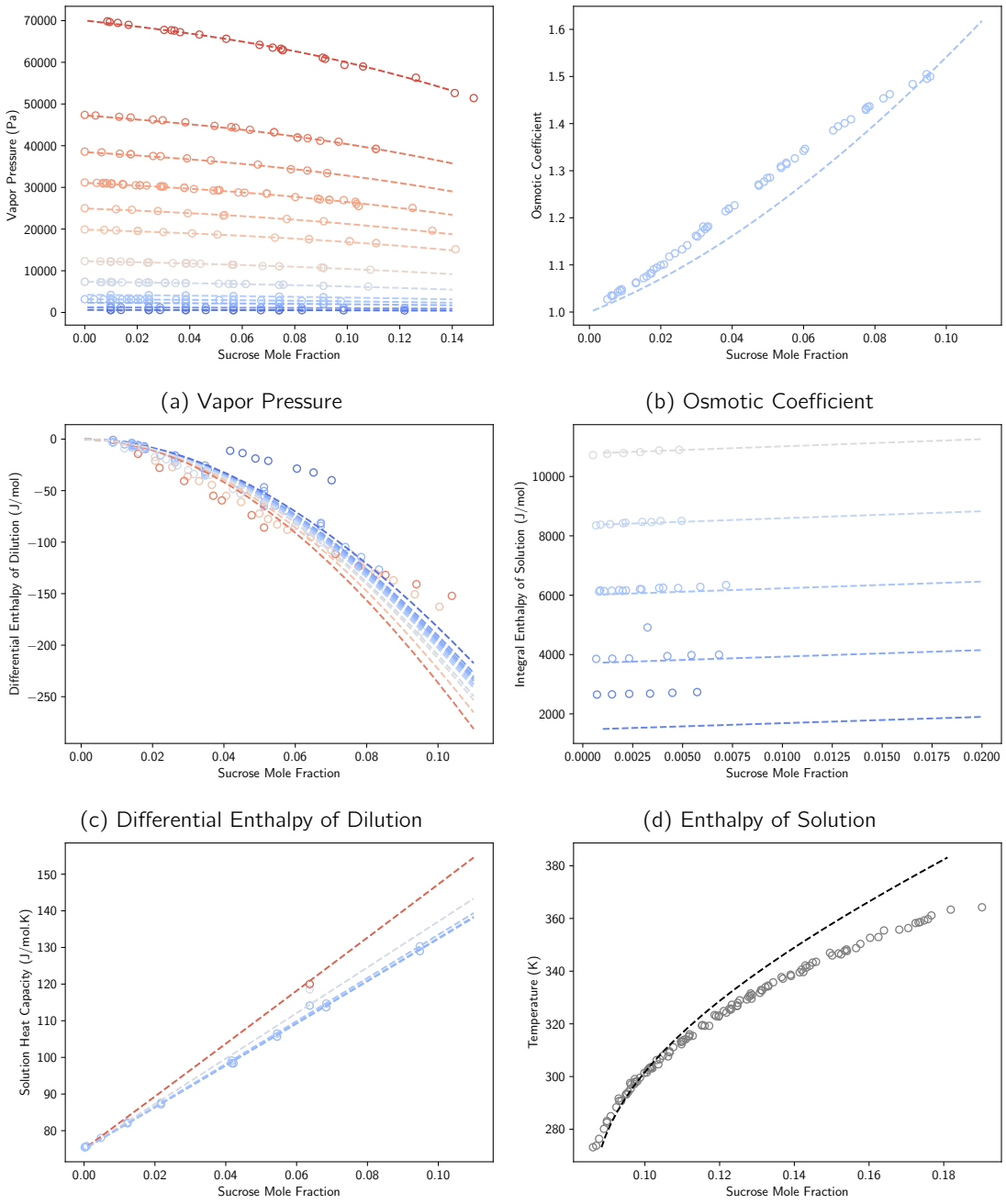
(c) Differential Enthalpy of Dilution

(d) Enthalpy of Solution

(e) Solution Heat Capacity

(f) Solubility

Figure 5.21: Model I. Open symbols correspond to experimental data (vapor pressure [168, 242, 164, 165, 172, 150, 245], Osmotic coefficient [170, 247], differential enthalpy of dilution [260, 176, 261], enthalpy of solution [259], solution heat capacity [178, 180] and solubility [147, 149, 150, 151]). Dashed lines correspond to predictions from global fit.



(a) Vapor Pressure

(b) Osmotic Coefficient

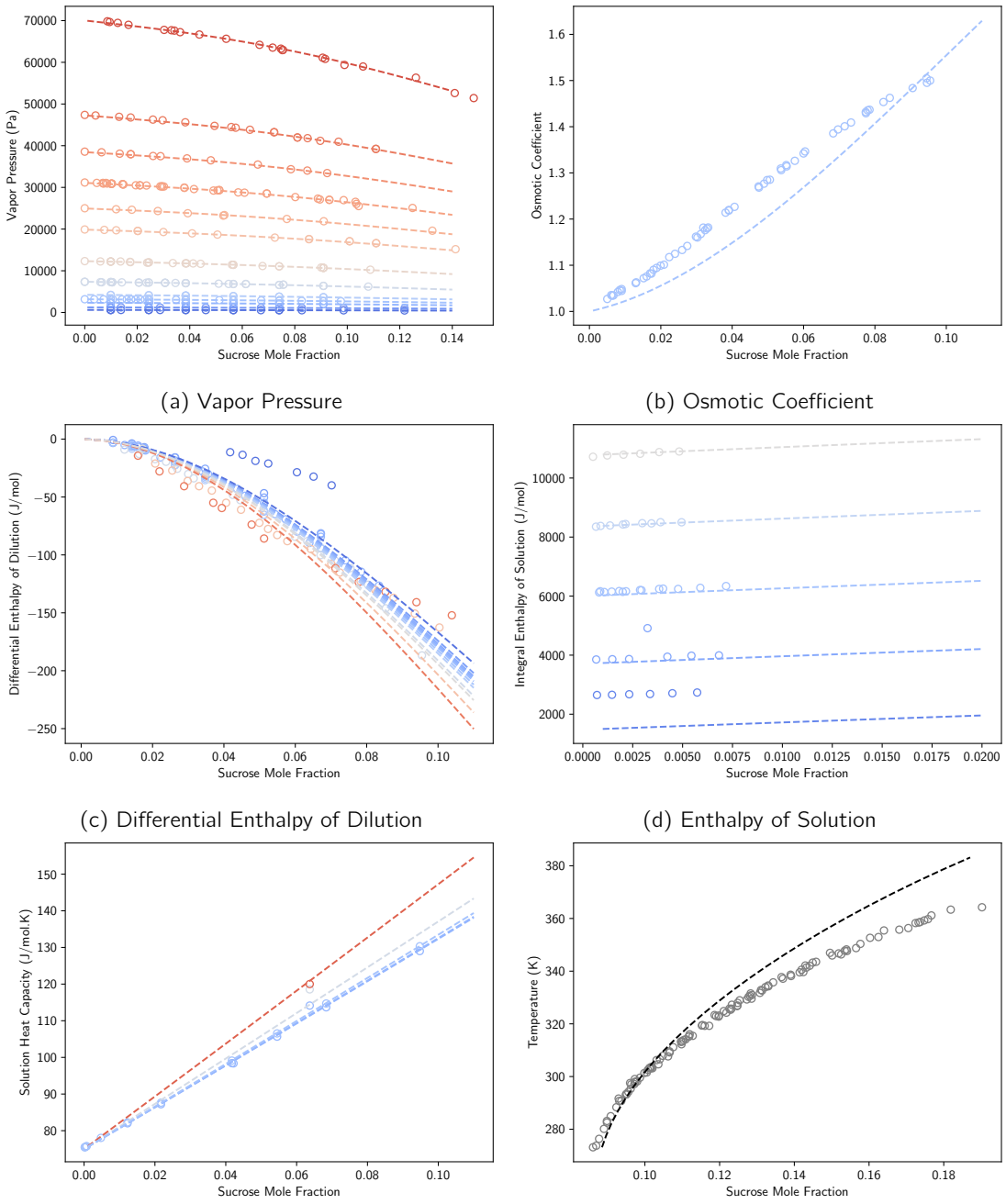
(c) Differential Enthalpy of Dilution

(d) Enthalpy of Solution

(e) Solution Heat Capacity

(f) Solubility

Figure 5.22: Model II. Open symbols correspond to experimental data (vapor pressure [168, 242, 164, 165, 172, 150, 245], Osmotic coefficient [170, 247], differential enthalpy of dilution [260, 176, 261], enthalpy of solution [259], solution heat capacity [178, 180] and solubility [147, 149, 150, 151]). Dashed lines correspond to predictions from global fit.



(a) Vapor Pressure

(b) Osmotic Coefficient

(c) Differential Enthalpy of Dilution

(d) Enthalpy of Solution

(e) Solution Heat Capacity

(f) Solubility

Figure 5.23: Model III. Open symbols correspond to experimental data (vapor pressure [168, 242, 164, 165, 172, 150, 245], Osmotic coefficient [170, 247], differential enthalpy of dilution [260, 176, 261], enthalpy of solution [259], solution heat capacity [178, 180] and solubility [147, 149, 150, 151]). Dashed lines correspond to predictions from global fit.

and 5.23, models II and III predict osmotic coefficients in poor agreement with those derived from experimental data, in agreement with the objective function contribution analysis described above.

Interestingly, none of the models explored were able to capture the temperature and composition dependence of the differential enthalpy of dilution data reported by [261]. More specifically, each model predicts the same functional dependence at all temperatures, which is in excellent agreement with experimental data reported at and around 298.15 K; however, is disagreement with measurements reported by [261] at 273.15, 333.15 and 353.15 K which, as discussed already, the validity of which is unclear due to a lack of independent supporting data reports. The need for additional enthalpy of dilution measurements was also highlighted in a recent work [264].

Finally, all models predict temperature dependent enthalpy of solution between 283.15–318.15 K at extremely dilute sucrose concentrations in excellent agreement with measurements reported in [259], noting again that only a single measure at 298.15 K was used during fitting. It should be noted that all models failed to capture data at 278.15 K.

Included in Fig. 5.21, 5.22 and 5.23 is a solubility prediction derived from the pure solid solute and respective multi-component liquid solution model compared to selected solubility data reported in the literature [147, 149, 150, 151] (sub-figure (f)). Each solubility curve was estimated by finding the sucrose-water liquid mixture composition that solves $\beta\mu_1^{s,\circ}(T) = \beta\mu_1^{l,\circ}(T)$ at fixed temperature, where $\mu_1^{s,\circ}(T)$ is the chemical potential of pure crystalline sucrose derived from the single component model developed earlier and $\mu_1^{l,\circ}(T)$ is the chemical potential of sucrose in the sucrose-water liquid mixture derived from the associated sucrose-water liquid mixture Gibbs free energy model. As an aside, the solubility calculation was completed using the Python modules described in Section 5.3.

From Fig. 5.21, model I gives a solubility prediction in excellent agreement both qualitatively and quantitatively with directly measured solubility data reported in the literature across the entire range of temperature for which data is available.

In contrast, from Fig. 5.22 and 5.23, models II and III are only able to predict sucrose solubility in qualitative agreement with directly measured data, diverging rapidly from the solubility at 298.15 K which was used a reference point. With reference to Chapter 3, the local slope of a solubility curve is defined by a ratio of the partial molar enthalpy of solution and the first partial derivative of the solute activity coefficient, both evaluated at saturation. Based on the aforementioned discrepancy between osmotic coefficients predicted by model II and III and available experimental data, this could explain the poor solubility performance, and highlights the potential challenges with predicting one set of thermodynamic properties based on mathematical modelling of other thermodynamic properties.

Reduced Dataset

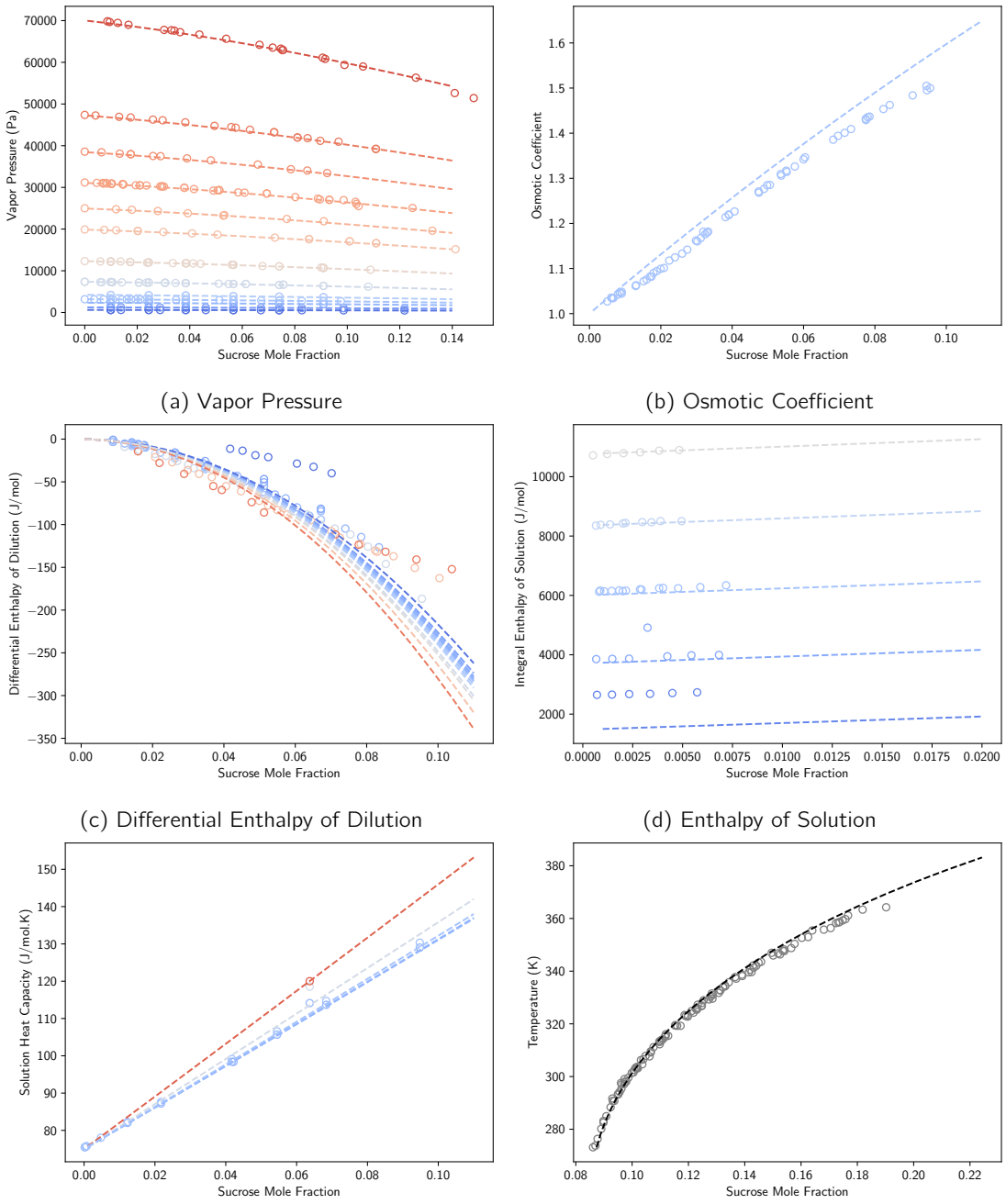
In the previous section, it was found that a solubility of sucrose is water was reliably predicted from the eutectic temperature to the pure vapor (solvent)-pure solid (solute)-liquid mixture (solute and solvent) triple point, with model I, potentially indicating that the 'real' sucrose-water liquid mixtures is well described by this model. As noted already, the form of this model is not flexible enough to capture the temperature and composition dependence suggested by differential enthalpy of dilution data at 273.15, 333.15 and 353.15 K reported in [261], and potentially To re-cap, as discussed in Section 5.5, there is significant uncertainty associated the temperature and composition of differential enthalpy of dilution data in general but, more specifically, with that reported in [261], given a lack of supporting data from independent reports. In this section, the parameters in model I are refined by re-fitting to available experimental data; however, differential enthalpy of dilution data at temperatures reported in [261] are excluded. The purpose of this is to provide a set of parameters, free from bias introduced by the inclusion of inconsistent data during the parameterisation process, which can be used in future modelling applications

The resulting best fit to sucrose-water solution properties is shown in Fig. 5.24. As shown, the resulting model predicts thermodynamic properties with to a comparable level to those already described for model I. The updated model parameters are provided in Table 5.5.

5.8 General Discussion

In this analysis, a straightforward and computationally convenient approach was used to parameterise each model, where an initial set of parameter values were specified and the Nelder-Mead numerical optimisation algorithm used. Assuming the chosen goodness-of-fit measure is appropriate, like with any optimisation problem, it is possible that the solutions found are sub-optimal due to the algorithm getting 'stuck' in local minima. Within the context of the methods applied here, this could be addressed by sampling different initial values and comparing the outputs. Alternatively, this could be addressed by exploring the application of more sophisticated numerical optimisation algorithms, such as 'basin hopping' [265], 'Differential Evolution (DE)' [266] or 'Particle Swarm Optimisation (PSO)' [267], all of which have been proposed to overcome potential challenges with practical optimisation problems; however, this is left as a recommendation for future work.

A key limitation of the analysis presented in this chapter is related to the estimation of only a single set of 'bets-fit' parameters for each of the sucrose-water liquid mixtures models examined. As illustrated in Chapter 3, it is possible to supplement a 'best fit' by perturbing the underlying experimental data within the estimated uncertainty and performing an independent parameterisation to this 'blurred' dataset, in an attempt to capture uncertainty introduced by experimental



(a) Vapor Pressure

(b) Osmotic Coefficient

(c) Differential Enthalpy of Dilution

(d) Enthalpy of Solution

(e) Solution Heat Capacity

(f) Solubility

Figure 5.24: Model I (reduced data set). Open symbols correspond to experimental data (vapor pressure [168, 242, 164, 165, 172, 150, 245], Osmotic coefficient [170, 247], differential enthalpy of dilution [260, 176, 261], enthalpy of solution [259], solution heat capacity [178, 180] and solubility [147, 149, 150, 151]). Dashed lines correspond to predictions from global fit.

data and modelling approach. Based on the computational approach applied, the time required to complete parameterisations was approximately 20, 40 and 60 minutes for model I, II and III, respectively. In Chapter 3, prediction bands were generated based on 1,000 perturbations of the underlying experimental data and subsequent re-parameterisation of the associated model. If a similar analysis was to be performed for the analysis described in this chapter, it calculations would be expected to take approximately 14, 28 and 60 days for model I, II and III, respectively (assuming that the computation time scales linearly with the number of parameterisations required). Of course, this could be reduced if the underlying code used to perform calculations was made more efficient.

There are limited (if any) studies focused on understanding how the characteristics (e.g., type, quality, distribution, etc.) of experimental data, coupled with applied modelling strategies, might impact the ability to estimate thermodynamic quantities in regions of the phase diagram where they have not yet been measured.

In a similar vein, as mentioned earlier in this chapter, recent work has focused on supplementing estimations of thermodynamic quantities and phase equilibria based on interpolations and extrapolations based on CALPHAD assessments. The challenges of this were partially addressed in a recent review [268] focused on uncertainty reduction and quantification in computational thermodynamics, including a discussion on identifying valuable measurements in the context of applying Bayesian methods for uncertainty quantification and experimental design during CALPHAD assessments. The author notes three challenges with attempting to quantify the 'value' of experiments in terms of reliably estimating quantities of interest (QoIs). The first challenge comes from the fact that the Gibbs free energy is not directly observable and so its uncertainty is, by definition, estimated by propagation from measureable quantities. The second challenge comes from the fact that it is only possible to specify Gibbs free energy models for a given system using a collection of several different experimental quantities. The final challenge is related to phase stability and our inability to know *a priori* what phases may exist in a given system.

In this chapter, Gibbs free energy models for single and multicomponent phases were constructed by correlating experimentally determined thermodynamic quantities, as widely applied in the literature. In Section 5.2.2, the relationship between a subset of often reported experimentally measureable quantities and the Gibbs free energy models applied in this work was discussed and the minimum requirements for model parameterisation outlined. In principle, when trying to construct Gibbs free energy models, it might be expected that different data types provide different information value when developing Gibbs free energy models. This general problem has been addressed as far back at 1978 in [269], where the authors attempted to gain insight into the types of data that would lead to reliable model structure and parameter estimates in various excess Gibbs free energy models (e.g., Van Laar, Wilson, UNIQUAC (and modifications), and Zeta) for the description of

both fully and partially miscible organic binary liquid mixtures. An interesting result from their analysis was that irrespective of the chosen excess Gibbs free energy model, being able to model one type of data (e.g., VLE) was not a sufficient criterion to reliably estimate another (e.g., enthalpy of mixing) and vice versa. However, the authors noted that measuring activity coefficients at infinite dilution at one temperature and one set enthalpy of dilution measurements appeared sufficient to parameterise all models investigated, giving reliable estimates of vapor pressure as a function of composition over a narrow range of temperature near those where the actual data was taken.

5.9 Conclusions

In this chapter, a general theoretical approach to the prediction of solubility as a function of temperature was described and it's application in the context of predicting the solubility of sucrose in water discussed. As shown, using a widely applied thermodynamic modelling framework (i.e., CALPHAD method), it is possible to predict solubility with only a single solubility measurement and no pure solute melting properties, in line with the general aim of this thesis. However, in contrast to the predictive approaches described in Chapters 3 and 4, the general approach presented in this chapter is not limited by the inclusion of approximations, in theory.

For the specific application to the sucrose-water system, we have shown that it is possible to predict solubility from the eutectic point to the pure crystal-liquid mixture-pure vapor triple point using only a single solubility estimate at a chosen reference temperature, taken here as 298.15 K, and a subset of experimentally measured thermodynamic data reported in the literature. Interestingly, this model predicted solution properties (i.e., differential enthalpy of dilution) in poor agreement with experimental data at elevated temperatures and high sucrose concentrations, which may be important given that this data has been used in several published reports focused on modelling the sucrose-water system. It was also shown that it was not possible to construct model that could simultaneously describe all available thermodynamic data for the sucrose-water liquid mixture using the Redlich-Kister polynomial with temperature dependent parameters commonly used in the CALPHAD community.

More generally, these results have implications for the work presented in Chapter 3. As discussed, the aqueous solubility of glycine polymorphs is subject to uncertainty with contradicting reports for α-glycine and very limited reports for β and γ. Given apparent challenges with reliably measuring glycine polymorph aqueous solubility across a broad range of temperature, a possible direction for future research would be to perform a systematic characterisation of glycine-water liquid mixtures and pure crystalline glycine polymorphs to supplement existing data and allow estimation in a similar manner to that illustrated for sucrose-water in this chapter. As mentioned

in [80], this would include additional measurements of pure crystal heat capacity beyond 310.15 K; however, should also include either enthalpy of solution and/or solution heat capacity measurements at temperatures up to 373.15 K from dilute to nearly saturated solutions. After which is should be possible to make predictions for glycine polymorph solubility, as well as being able to discriminate between different reports in literature in a more robust manner than conclusions drawn from approximate modeling in this thesis.

Chapter 6

Conclusions and Future Work

As discussed extensively throughout this thesis, accurate and reliable estimates of solubility are crucial for the design, development and optimisation of efficient crystallisation unit operations, where crystallisation is a separation and purification technique used ubiquitously throughout the chemical manufacturing industry. As such, this thesis has focussed on the general problem of developing and evaluating methods to allow estimation of the solubility of solids in liquids as a function of temperature at constant pressure—all of which are rooted in classical equilibrium thermodynamics. More specifically, we have proposed and evaluated new methods that reduce reliance on directly measured solubility, while simultaneously addressing potential limitations with already existing model-based approaches, such as reliance on pure solute melting properties which can be challenging to determine accurately.

In the remainder of this chapter, the main findings, contributions, and limitations of the work presented in this thesis are discussed. This is followed by recommendations for future work.

6.1 Conclusions

In Chapter 3 and 4, two methods were proposed that allowed prediction of solubility through estimation of thermodynamic quantities at a single point on the solubility curve. The first approach (i.e., Chapter 3), was based on the van't Hoff equation applied to solid-liquid equilibrium—which provides a mathematical relationship describing the rate of change of a solubility curve with respect to temperature in terms of a ratio between two thermodynamic quantities, both of which would be expected to vary with temperature and composition. By interpreting the temperature and composition dependence of the numerator and denominator using a 1^{st}-order Taylor series, it was possible to numerically integrate the approximate model and predict solubility, as long as all terms introduced by the expansion were evaluated. Thermodynamic quantities required to make predictions were estimated for the α-glycine-water system via mathematical modeling of

experimental data reported in the literature at 298.15 K, and results in good agreement with directly measured solubility from 260–340 K were shown.

The second approach (i.e., Chapter 4), took learnings from work presented in Chapter 3, proposing an analytical five parameter model describing the temperature dependence of solubility by interpreting the difference in pure crystal and dissolved solute chemical potential (at saturation), in terms of a 2^{nd}-order Taylor series expansion. As with the approach presented in Chapter 3, the model proposed in Chapter 4 made it possible to predict solubility as function of temperature using only a single solubility estimate, as long as estimates of additional thermodynamic quantities evaluated at the same point on the solubility curve were known. The predictive ability of this approach was examined by estimating the aqueous solubility of three industrially relevant organic solutes (i.e., α-glycine, sucrose and urea), and similar results were found when compared to Chapter 3—when reliable experimental data is available near the solubility point taken as a reference, predictions in good agreement with directly measured solubility within ± 40 K of 298.15 K.

In Chapter 5, a more general approach to the estimation of solubility was reviewed, where Gibbs free energy functions describing all phases of a hypothetical system were modeled explicitly. In contrast to Chapter 3 and 4, this approach wasn't restricted to the estimation of thermodynamic quantities at a single point on the solubility curve, more rather any (and all) valid experimental data was used to parameterise the underling Gibbs free energy models required to estimate solubility. However, it was shown for a set of widely applied Gibbs free energy models that it is theoretically possible to completely parameterise all models required to estimate solubility across an unlimited range of temperature using the same types of experimental data used in Chapter 3 and 4, as long as only a single measurement of solubility is known. This approach was illustrated for the sucrose-water system, which has a significant amount of thermodynamic data across a broad range of temperature coupled with well defined phase boundaries, and a simple model was found that could completely describe the solubility curve from the eutectic point up to the crystalline solute-solution-vapor triple point.

The main contribution of the work presented in this thesis are new methods that allow estimation of the solubility of solids in liquids as a function of temperature. Crucially; however, all approaches make it possible to do so without knowledge of the pure solute melting properties and with only a single measure of solubility, which can be viewed as an advance on the current state of the art. Comparing the approaches in Chapter 3 and 4 to that in Chapter 5, the former provide estimates in the local vicinity of the chosen solubility point, while the latter can provide estimates across an unbounded range of temperature, if reliable data other than solubility is available and mathematical modelling is applied. The work presented in this thesis should be of interest to practitioners with an interest in estimating the solubility solids in liquids, where conventional approaches (both experimental and model-based) may not yield desired accuracy. For example,

compounds like β-glycine (see Chapter 3) where measurement of solubility is challenging at any temperature but measurement of other thermodynamic quantities involving the pure crystal (i.e., crystalline heat capacity, enthalpy of solution) and solution (i.e., solution heat capacity, vapor pressure, osmotic coefficient, etc.) might be easier and more reliable. This is directly applicable to the pharmaceutical industry, where novel solid compounds are produced with characteristics that make direct measurement of solubility challenging, when compared to the simple model systems used throughout this thesis.

It is important to highlight the primary challenge with the methods explored in this thesis, namely the reliance on mathematical modeling as a tool used to translate thermodynamic information from measurable properties other than solubility, into an estimate for solubility. As illustrated throughout this thesis, subjective choices made during mathematical modeling (e.g., measure of best fit, optimisation algorithm, model and data), can have a meaningful impact on the approaches. However, uncertainty introduced as a result of applying mathematical modeling techniques can be reduced by ensuring that the data used to make predictions is reliable, numerous and covers an appropriate range conditions—something that is challenging to know for a given application.

6.2 Future Work

Perhaps the most obvious direction for future research would be the expansion of described techniques to a broader space of chemical compounds. In this work, analysis is focused on a limited set of organic compounds (i.e., glycine, urea and sucrose) and their solubility in water. As alluded to already, this was driven by necessity rather than by choice. As mentioned briefly throughout this thesis, these systems have a reasonable amount of thermodynamic data reported in the literature covering both single component and multi-component phases, as well as and phase equilibria over broad ranges of temperature and composition, when compared to many other systems—obvious counter examples being salt-water systems (e.g., sodium chloride-water and potassium-chloride-water). While water is an important solvent, and glycine, urea and sucrose have somewhat diverse chemical characteristics (i.e., functional groups, solubility behavior with respect to temperature, molecular weight, etc.,), they represent an extremely small set of possible chemical properties, relative to the broader universe of chemical compounds encountered in chemical manufacturing processes.

Furthermore, all analysis presented in this thesis is focused on binary systems (i.e., single solute, single solvent). Despite being an import class of crystallisation system, as described in Chapter 2, industrial crystallisation is often achieved with more complex systems, an example being anti-solvent crystallisation, where multiple solvents are present. From a theoretical per-

spective, extension of the results approaches described in this thesis to more complex systems is possible; however, this is currently limited by limited availability of comprehensive thermodynamic characterisation of ternary (and higher) systems.

Again, work presented in this thesis was driven by availability of experimental data. However, thermodynamic data is usually collected with a specific outcome in mind which is usually not for the subsequent prediction of solubility of solids in liquids. From a practical perspective, it would be advantageous to have coordinated efforts to produce comprehensive thermodynamic assessments of industrially relevant crystallisation systems. As shown in Chapter 5, a minimum amount of information is required to develop reliable estimates for the phase diagram of binary systems, focused on the region bounding the assumed solute-solvent liquid mixture, is a single solubility estimate, a single enthalpy of solution measurement and a complete determination of the solution heat capacity across the entire range of solution temperature and composition.

In Chapters 3 and 4, approximate models were proposed allowing prediction of solubility based on thermodynamic quantities evaluated at an chosen point on a given solubility curve. For the model systems explored in this thesis, estimating these quantities using available experimental data involved a multi-step processes involving development of mathematical models, from which quantities were evaluated, introducing undefined uncertainty. A possible direction for future work would be the development and assessment of novel experimental techniques allowing direct measurement of thermodynamic quantities. For example, it has been shown that dynamic light scattering can be used to measure the first composition partial derivative (i.e., $\partial \ln a_i / \partial \ln x_i$) of components in liquid mixtures. Furthermore, terms related to the difference in partial molar enthalpy can be estimated from a hypothetical process defined as the differential enthalpy of solution (c.f. differential enthalpy of dilution). In principle, it should be possible to make these measurements; however, as far as we are aware, they are yet to be reported in the literature.

Finally, well known functional forms for the Gibbs free energy of single and multi-component phases were used throughout this thesis (e.g., Scatchard-Hildebrand(-Flory-Huggins), Redlich-Kister, etc) for mathematical modeling purposes. Another direction for future research—particularly with reference to results presented in Chapter 5—could be to understand the applicability of non-traditional (in the context of thermodynamic modeling) modeling techniques for the representation of Gibbs free energies. More specifically, this may include neural networks, noting that recent work has shown that a single layer neural network was able to represent the Gibbs free energy of pure iron between 0 and 6,000 K [270].

Bibliography

[1] Richard Turton et al. *Analysis, synthesis, and design of chemical processes*. 5th ed. International Series in the Physical and Chemical Engineering Sciences. Upper Saddle River, NJ: Pearson, Oct. 2020.

[2] Don W Green and Marylee Z Southard. *Perry's chemical engineers' handbook, 9th edition*. 9th ed. McGraw-Hill Education, July 2018.

[3] Wolfgang Beckmann, ed. *Crystallization: basic concepts and industrial applications*. eng. Weinheim: Wiley-VCH, 2013. ISBN: 9783527650354.

[4] Deniz Erdemir, Alfred Y Lee, and Allan S Myerson. "Nucleation of crystals from solution: classical and two-step models". en. In: *Acc. Chem. Res.* 42.5 (May 2009), pp. 621–629.

[5] Herbert B Callen. *Thermodynamics and an Introduction to Thermostatistics*. en. 2nd ed. Nashville, TN: John Wiley & Sons, Aug. 1985.

[6] J. W. Mullin. *Crystallization*. 4th ed. Oxford ; Boston: Butterworth-Heinemann, 2001. ISBN: 9780750648332.

[7] Allan S Myerson. *Handbook of industrial crystallization*. 2nd ed. Woburn, MA: Butterworth-Heinemann, Jan. 2002.

[8] Alison Lewis et al. *Industrial Crystallization*. Cambridge University Press, 2015. DOI: 10.1017/cbo9781107280427. URL: https://doi.org/10.1017/cbo9781107280427.

[9] Govindhan Dhanaraj et al., eds. *Springer handbook of crystal growth*. en. 2010th ed. Springer Handbooks. Berlin, Germany: Springer, Oct. 20.

[10] Shijie Xu et al. "Overview of Secondary Nucleation: From Fundamentals to Application". In: *Industrial & Engineering Chemistry Research* 59.41 (Sept. 2020), pp. 18335–18356. DOI: 10.1021/acs.iecr.0c03304. URL: https://doi.org/10.1021/acs.iecr.0c03304.

[11] António M. Peres and Eugénia A. Macedo. "A modified UNIFAC model for the calculation of thermodynamic properties of aqueous and non-aqueous solutions containing sugars". In: *Fluid Phase Equilibria* 139.1-2 (Dec. 1997), pp. 47–74. DOI: 10.1016/s0378-3812(97)00196-9. URL: https://doi.org/10.1016/s0378-3812(97)00196-9.

[12] Kyungho Park, James M. B. Evans, and Allan S. Myerson. "Determination of Solubility of Polymorphs Using Differential Scanning Calorimetry". In: *Crystal Growth & Design* 3.6 (Sept. 2003), pp. 991–995. DOI: 10.1021/cg0340502. URL: https://doi.org/10.1021/cg0340502.

[13] Dóra Csicsák et al. "Towards more accurate solubility measurements with real time monitoring: a carvedilol case study". In: *New Journal of Chemistry* 45.26 (2021), pp. 11618–11625. DOI: 10.1039/d1nj01349a. URL: https://doi.org/10.1039/d1nj01349a.

[14] Frans L. Muller, Mark Fielding, and Simon Black. "A Practical Approach for Using Solubility to Design Cooling Crystallisations". In: *Organic Process Research & Development* 13.6 (Oct. 2009), pp. 1315–1321. DOI: 10.1021/op9001438. URL: https://doi.org/10.1021/op9001438.

[15] Simon Black et al. "On the Measurement of Solubility". In: *Organic Process Research &; Development* 17.3 (Feb. 2013), pp. 486–492. DOI: 10.1021/op300336n. URL: https://doi.org/10.1021/op300336n.

[16] Somnath S. Kadam et al. "A new view on the metastable zone width during cooling crystallization". In: *Chemical Engineering Science* 72 (Apr. 2012), pp. 10–19. DOI: 10.1016/j.ces.2012.01.002. URL: https://doi.org/10.1016/j.ces.2012.01.002.

[17] Peter Rudolph, Thomas F. Kuech, and Tatau Nishinga. *Handbook of crystal growth*. eng. 2nd rev. ed. Amsterdam: Elsevier, 2015. ISBN: 9780444563699 9780444633033 9780444633040.

[18] Jr. Howard K. Zimmerman. "The Experimental Determination of Solubilities." In: *Chemical Reviews* 51.1 (Aug. 1952), pp. 25–65. DOI: 10.1021/cr60158a002. URL: https://doi.org/10.1021/cr60158a002.

[19] G.T. Hefter and R.P.T. Tomkins, eds. *The Experimental Determination of Solubilities*. John Wiley & Sons, Ltd, Mar. 2003. DOI: 10.1002/0470867833. URL: https://doi.org/10.1002/0470867833.

[20] Erich Königsberger. "Editorial: Guidelines for the Measurement of Solid–Liquid Solubility Data at Atmospheric Pressure". In: *Journal of Chemical & Engineering Data* 64.2 (Feb. 2019), pp. 381–385. DOI: 10.1021/acs.jced.8b01263. URL: https://doi.org/10.1021/acs.jced.8b01263.

[21] Rajeev Mohan, Heike Lorenz, and Allan S. Myerson. "Solubility Measurement Using Differential Scanning Calorimetry". In: *Industrial & Engineering Chemistry Research* 41.19 (Aug. 2002), pp. 4854–4862. DOI: 10.1021/ie0200353. URL: https://doi.org/10.1021/ie0200353.

[22] Edit Baka, John E.A. Comer, and Krisztina Takács-Novák. "Study of equilibrium solubility measurement by saturation shake-flask method using hydrochlorothiazide as model compound". In: *Journal of Pharmaceutical and Biomedical Analysis* 46.2 (Jan. 2008), pp. 335–341. DOI: 10.1016/j.jpba.2007.10.030. URL: https://doi.org/10.1016/j.jpba.2007.10.030.

[23] Andrew Cashmore. *Understanding, measurement and control of secondary nucleation*. 2022. URL: Uhttps://doi.org/10.48730/p8cj-xq09.

[24] Lai Zeng, Åke C. Rasmuson, and Michael Svärd. "Solubility of Two Polymorphs of Tolbutamide in n-Propanol: Comparison of Methods". In: *Journal of Pharmaceutical Sciences* 109.10 (Oct. 2020), pp. 3021–3026. DOI: 10.1016/j.xphs.2020.06.022. URL: https://doi.org/10.1016/j.xphs.2020.06.022.

[25] Yongjin Yi et al. "Development of a Small-Scale Automated Solubility Measurement Apparatus". In: *Industrial & Engineering Chemistry Research* 44.15 (June 2005), pp. 5427–5433. DOI: 10.1021/ie049215y. URL: https://doi.org/10.1021/ie049215y.

[26] Fabian B. Thygs, Juliane Merz, and Gerhard Schembecker. "Automation of Solubility Measurements on a Robotic Platform". In: *Chemical Engineering & Technology* 39.6 (Mar. 2016), pp. 1049–1057. DOI: 10.1002/ceat.201500572. URL: https://doi.org/10.1002/ceat.201500572.

[27] Parisa Shiri et al. "Automated solubility screening platform using computer vision". In: *iScience* 24.3 (Mar. 2021), p. 102176. DOI: 10.1016/j.isci.2021.102176. URL: https://doi.org/10.1016/j.isci.2021.102176.

[28] Fredrik L Nordström and Åke C Rasmuson. "Determination of the activity of a molecular solute in saturated solution". en. In: *J. Chem. Thermodyn.* 40.12 (Dec. 2008), pp. 1684–1692.

[29] Alexander P. Bünz, Björn Braun, and Ralf Janowsky. "Application of Quantitative Structure-Performance Relationship and Neural Network Models for the Prediction of Physical Properties from Molecular Structure". In: *Industrial & Engineering Chemistry Research* 37.8 (May 1998), pp. 3043–3051. DOI: 10.1021/ie970910y. URL: https://doi.org/10.1021/ie970910y.

[30] Martin Kuentz and Christel A.S. Bergström. "Synergistic Computational Modeling Approaches as Team Players in the Game of Solubility Predictions". In: *Journal of Pharmaceutical Sciences* 110.1 (Jan. 2021), pp. 22–34. DOI: 10.1016/j.xphs.2020.10.068. URL: https://doi.org/10.1016/j.xphs.2020.10.068.

[31] A. Shayanfar, M.A.A. Fakhree, and A. Jouyban. "A simple QSPR model to predict aqueous solubility of drugs". In: *Journal of Drug Delivery Science and Technology* 20.6 (2010), pp. 467–476. DOI: 10.1016/s1773-2247(10)50080-7. URL: https://doi.org/10.1016/s1773-2247(10)50080-7.

[32] Yingqing Ran, Neera Jain, and Samuel H. Yalkowsky. "Prediction of Aqueous Solubility of Organic Compounds by the General Solubility Equation (GSE)". In: *Journal of Chemical Information and Computer Sciences* 41.5 (July 2001), pp. 1208–1217. DOI: 10.1021/ci010287z. URL: https://doi.org/10.1021/ci010287z.

[33] Michael H. Abraham and Joelle Le. "The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship". In: *Journal of Pharmaceutical Sciences* 88.9 (Sept. 1999), pp. 868–880. DOI: 10.1021/js9901007. URL: https://doi.org/10.1021/js9901007.

[34] Zhuyifan Ye and Defang Ouyang. "Prediction of small-molecule compound solubility in organic solvents by machine learning algorithms". In: *Journal of Cheminformatics* 13.1 (Dec. 2021). DOI: 10.1186/s13321-021-00575-3. URL: https://doi.org/10.1186/s13321-021-00575-3.

[35] J. M Smith et al. *Introduction to chemical engineering thermodynamics*. English. OCLC: 1242912106. 2022. ISBN: 9781264364596.

[36] J. M. Prausnitz, Ruediger N. Lichtenthaler, and Edmundo Gomes de Azevedo. *Molecular thermodynamics of fluid-phase equilibria*. 2nd ed. Englewood Cliffs, N.J: Prentice-Hall, 1986. ISBN: 9780135995648.

[37] J. J. van Laar. "Über Dampfspannungen von binären Gemischen". In: *Zeitschrift für Physikalische Chemie* 72U.1 (Mar. 1910), pp. 723–751. DOI: 10.1515/zpch-1910-7236. URL: https://doi.org/10.1515/zpch-1910-7236.

[38] Joel H. Hildebrand. "SOLUBILITY. XII. REGULAR SOLUTIONS". In: *Journal of the American Chemical Society* 51.1 (Jan. 1929), pp. 66–80. DOI: 10.1021/ja01376a009. URL: https://doi.org/10.1021/ja01376a009.

[39] George. Scatchard. "Equilibria in Non-electrolyte Solutions in Relation to the Vapor Pressures and Densities of the Components." In: *Chemical Reviews* 8.2 (Apr. 1931), pp. 321–333. DOI: 10.1021/cr60030a010. URL: https://doi.org/10.1021/cr60030a010.

[40] Otto Redlich and A. T. Kister. "Algebraic Representation of Thermodynamic Properties and the Classification of Solutions". In: *Industrial & Engineering Chemistry* 40.2 (Feb. 1948), pp. 345–348. DOI: 10.1021/ie50458a036. URL: https://doi.org/10.1021/ie50458a036.

[41] Grant M. Wilson. "Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing". In: *Journal of the American Chemical Society* 86.2 (Jan. 1964), pp. 127–130. DOI: 10.1021/ja01056a002. URL: https://doi.org/10.1021/ja01056a002.

[42] Henri Renon and J. M. Prausnitz. "Local compositions in thermodynamic excess functions for liquid mixtures". In: *AIChE Journal* 14.1 (Jan. 1968), pp. 135–144. DOI: 10.1002/aic.690140124. URL: https://doi.org/10.1002/aic.690140124.

[43] Denis S. Abrams and John M. Prausnitz. "Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems". In: *AIChE Journal* 21.1 (Jan. 1975), pp. 116–128. DOI: 10.1002/aic.690210115. URL: https://doi.org/10.1002/aic.690210115.

[44] Jorge Marrero and Rafiqul Gani. "Group-contribution based estimation of pure component properties". In: *Fluid Phase Equilibria* 183-184 (July 2001), pp. 183–208. DOI: 10.1016/s0378-3812(01)00431-9. URL: https://doi.org/10.1016/s0378-3812(01)00431-9.

[45] Jürgen Gmehling, Dana Constantinescu, and Bastian Schmid. "Group Contribution Methods for Phase Equilibrium Calculations". In: *Annual Review of Chemical and Biomolecular Engineering* 6.1 (July 2015), pp. 267–292. DOI: 10.1146/annurev-chembioeng-061114-123424. URL: https://doi.org/10.1146/annurev-chembioeng-061114-123424.

[46] Aage Fredenslund, Russell L. Jones, and John M. Prausnitz. "Group-contribution estimation of activity coefficients in nonideal liquid mixtures". In: *AIChE Journal* 21.6 (Nov. 1975), pp. 1086–1099. DOI: 10.1002/aic.690210607. URL: https://doi.org/10.1002/aic.690210607.

[47] Ulrich Weidlich and Juergen Gmehling. "A modified UNIFAC model. 1. Prediction of VLE, hE, and .gamma..infin." In: *Industrial & Engineering Chemistry Research* 26.7 (July 1987), pp. 1372–1381. DOI: 10.1021/ie00067a018. URL: https://doi.org/10.1021/ie00067a018.

[48] Anja Diedrichs and Jürgen Gmehling. "Solubility Calculation of Active Pharmaceutical Ingredients in Alkanes, Alcohols, Water and their Mixtures Using Various Activity Coefficient Models". In: *Industrial & Engineering Chemistry Research* 50.3 (Dec. 2010), pp. 1757–1769. DOI: 10.1021/ie101373k. URL: https://doi.org/10.1021/ie101373k.

[49] Nikolaos Spiliotis and Dimitrios Tassios. "A UNIFAC model for phase equilibrium calculations in aqueous and nonaqueous sugar solutions". en. In: *Fluid Phase Equilib.* 173.1 (Aug. 2000), pp. 39–55.

[50] Chau-Chyun Chen and Yuhua Song. "Solubility Modeling with a Nonrandom Two-Liquid Segment Activity Coefficient Model". In: *Industrial & Engineering Chemistry Research* 43.26 (Dec. 2004), pp. 8354–8362. DOI: 10.1021/ie049463u. URL: https://doi.org/10.1021/ie049463u.

[51] Ali Haghtalab and Jaber Yousefi Seyf. "Vapor–Liquid and Solid–Liquid Modeling with a Universal Quasichemical Segment-Based Activity Coefficient Model". In: *Industrial & Engineering Chemistry Research* 54.34 (Aug. 2015), pp. 8611–8623. DOI: 10.1021/acs.iecr.5b01573. URL: https://doi.org/10.1021/acs.iecr.5b01573.

[52] Benedikt Winter et al. "SPT-NRTL: A physics-guided machine learning model to predict thermodynamically consistent activity coefficients". In: *Fluid Phase Equilibria* 568 (May 2023), p. 113731. DOI: 10.1016/j.fluid.2023.113731. URL: https://doi.org/10.1016/j.fluid.2023.113731.

[53] Vipul Mann, Rafiqul Gani, and Venkat Venkatasubramanian. "Group contribution-based property modeling for chemical product design: A perspective in the AI era". In: *Fluid Phase Equilibria* 568 (May 2023), p. 113734. DOI: 10.1016/j.fluid.2023.113734. URL: https://doi.org/10.1016/j.fluid.2023.113734.

[54] Andreas Klamt. "Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena". In: *The Journal of Physical Chemistry* 99.7 (Feb. 1995), pp. 2224–2235. DOI: 10.1021/j100007a062. URL: https://doi.org/10.1021/j100007a062.

[55] Shiang-Tai Lin and Stanley I. Sandler. "A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model". In: *Industrial & Engineering Chemistry Research* 41.5 (Dec. 2001), pp. 899–913. DOI: 10.1021/ie001047w. URL: https://doi.org/10.1021/ie001047w.

[56] Martin Klajmon. "Purely Predicting the Pharmaceutical Solubility: What to Expect from PC-SAFT and COSMO-RS?" In: *Molecular Pharmaceutics* 19.11 (Sept. 2022), pp. 4212–4232. DOI: 10.1021/acs.molpharmaceut.2c00573. URL: https://doi.org/10.1021/acs.molpharmaceut.2c00573.

[57] Svava Ósk Jónsdóttir and Peter Rasmussen. "Phase equilibria of carbohydrates in polar solvents". en. In: *Fluid Phase Equilib.* 158-160 (June 1999), pp. 411–418.

[58] Steven H. Neau, Satej V. Bhandarkar, and Eckhard W. Hellmuth. "Differential molar heat capacities to test ideal solubility estimations". In: *Pharmaceutical Research* 14.5 (1997), pp. 601–605. DOI: 10.1023/a:1012148910975. URL: https://doi.org/10.1023/a:1012148910975.

[59] S.H. Yalkowsky, G.L. Flynn, and G.L. Amidon. "Solubility of Nonelectrolytes in Polar Solvents". In: *Journal of Pharmaceutical Sciences* 61.6 (June 1972), pp. 983–984. DOI: 10.1002/jps.2600610643. URL: https://doi.org/10.1002/jps.2600610643.

[60] Steven H. Neau, Gordon L. Flynn, and Samuel H. Yalkowsky. "The influence of heat capacity assumptions on the estimation of solubility parameters from solubility data". In: *International Journal of Pharmaceutics* 49.3 (Feb. 1989), pp. 223–229. DOI: 10.1016/0378-5173(89)90346-3. URL: https://doi.org/10.1016/0378-5173(89)90346-3.

[61] Steven H. Neau and Gordon L. Flynn. In: *Pharmaceutical Research* 07.11 (1990), pp. 1157–1162. DOI: 10.1023/a:1015984310068. URL: https://doi.org/10.1023/a:1015984310068.

[62] Samuel H. Yalkowsky and Min Wu. "Estimation of the ideal solubility (crystal–liquid fugacity ratio) of organic compounds". In: *Journal of Pharmaceutical Sciences* 99.3 (Mar. 2010), pp. 1100–1106. DOI: 10.1002/jps.21897. URL: https://doi.org/10.1002/jps.21897.

[63] Vojtěch Štejfa et al. "Heat capacities of selected active pharmaceutical ingredients". In: *The Journal of Chemical Thermodynamics* 163 (Dec. 2021), p. 106585. DOI: 10.1016/j.jct.2021.106585. URL: https://doi.org/10.1016/j.jct.2021.106585.

[64] Mikhail I. Yagofarov and Boris N. Solomonov. "Interpolation of the Temperature Dependence of the Fusion Enthalpy of Aromatic Compounds Between 298.15 K and the Melting Temperature". In: *International Journal of Thermophysics* 43.6 (Apr. 2022). DOI: 10.1007/s10765-022-03020-1. URL: https://doi.org/10.1007/s10765-022-03020-1.

[65] Ingrid M. Weiss et al. "Thermal decomposition of the amino acids glycine, cysteine, aspartic acid, asparagine, glutamic acid, glutamine, arginine and histidine". In: *BMC Biophysics* 11.1 (Feb. 2018). DOI: 10.1186/s13628-018-0042-4. URL: https://doi.org/10.1186/s13628-018-0042-4.

[66] Yeong Zen Chua et al. "New experimental melting properties as access for predicting amino-acid solubility". In: *RSC Advances* 8.12 (2018), pp. 6365–6372. DOI: 10.1039/c8ra00334c. URL: https://doi.org/10.1039/c8ra00334c.

[67] Yrjö H. Roos et al. "Comment on the Melting and Decomposition of Sugars". In: *Journal of Agricultural and Food Chemistry* 60.41 (Oct. 2012), pp. 10359–10362. DOI: 10.1021/jf3002526. URL: https://doi.org/10.1021/jf3002526.

[68] Jürgen G. Gmehling, Thomas F. Anderson, and John M. Prausnitz. "Solid-Liquid Equilibria Using UNIFAC". In: *Industrial & Engineering Chemistry Fundamentals* 17.4 (Nov. 1978), pp. 269–273. DOI: 10.1021/i160068a008. URL: https://doi.org/10.1021/i160068a008.

[69] Sandra Gracin, Tore Brinck, and Åke C. Rasmuson. "Prediction of Solubility of Solid Organic Compounds in Solvents by UNIFAC". In: *Industrial & Engineering Chemistry Research* 41.20 (Sept. 2002), pp. 5114–5124. DOI: 10.1021/ie011014w. URL: https://doi.org/10.1021/ie011014w.

[70] Baptiste Bouillot, Sébastien Teychené, and Béatrice Biscans. "An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents". In: *Fluid Phase Equilibria* 309.1 (Oct. 2011), pp. 36–52. DOI: 10.1016/j.fluid.2011.06.032. URL: https://doi.org/10.1016/j.fluid.2011.06.032.

[71] Sara A. Febra et al. "Extending the SAFT-γ Mie approach to model benzoic acid, diphenylamine, and mefenamic acid: Solubility prediction and experimental measurement". In: *Fluid Phase Equilibria* 540 (July 2021), p. 113002. DOI: 10.1016/j.fluid.2021.113002. URL: https://doi.org/10.1016/j.fluid.2021.113002.

[72] R. Gary Hollenbeck. "Determination of Differential Heat of Solution in Real Solutions From Variation in Solubility With Temperature". In: *J. Pharm. Sci.* 69.10 (1980), pp. 1241–1242. DOI: 10.1002/jps.2600691038. URL: https://doi.org/10.1002/jps.2600691038.

[73] D.J.W. Grant et al. "Non-Linear Van't Hoff Solubility-Temperature Plots and Their Pharmaceutical Interpretation". In: *Int. J. Pharm.* 18.1-2 (1984), pp. 25–38. DOI: 10.1016/0378-5173(84)90104-2. URL: https://doi.org/10.1016/0378-5173(84)90104-2.

[74] A. T. Williamson. "The Exact Calculation of Heats of Solution From Solubility Data". In: *Trans. Faraday Soc.* 40.nil (1944), p. 421. DOI: 10.1039/tf9444000421. URL: https://doi.org/10.1039/tf9444000421.

[75] Fredrik L. Nordström and Åke C. Rasmuson. "Phase Equilibria and Thermodynamics of P-Hydroxybenzoic Acid". In: *J. Pharm. Sci.* 95.4 (2006), pp. 748–760. DOI: 10.1002/jps.20569. URL: https://doi.org/10.1002/jps.20569.

[76] Rodrigo Soto et al. "Ketoprofen Solubility in Pure Organic Solvents Using in Situ Ftir and Uv-Vis and Analysis of Solution Thermodynamics". In: *Org. Process Res. & Dev.* 25.11

(2021), pp. 2403–2414. DOI: 10.1021/acs.oprd.1c00156. URL: https://doi.org/10.1021/acs.oprd.1c00156.

[77] Charles A. Zittle and Carl L. A. Schmidt. "Heats of solution, heats of dilution, and specific heats of aqueous solutions of certain amino acids". In: *J. Biol. Chem.* 108 (1935), pp. 161–185.

[78] S.J Gill, N.F Nichols, and I Wadsö. "Calorimetric Determination of Enthalpies of Solution of Slightly Soluble Liquids Ii. Enthalpy of Solution of Some Hydrocarbons in Water and Their Use in Establishing the Temperature Dependence of Their Solubilities". In: *J. Chem. Thermodyn.* 8.5 (1976), pp. 445–452. DOI: 10.1016/0021-9614(76)90065-3. URL: https://doi.org/10.1016/0021-9614(76)90065-3.

[79] Benjamin D. Hamilton, Marc A. Hillmyer, and Michael D. Ward. "Glycine Polymorphism in Nanoscale Crystallization Chambers". In: *Cryst. Growth Des.* 8.9 (2008), pp. 3368–3375. DOI: 10.1021/cg800326a. URL: https://doi.org/10.1021/cg800326a.

[80] Darren Rowland. "Thermodynamic Properties of the Glycine + H2O System". In: *Journal of Physical and Chemical Reference Data* 47.2 (June 2018), p. 023104. DOI: 10.1063/1.5016677. URL: https://doi.org/10.1063/1.5016677.

[81] Maria J. Vesga et al. "Conundrum of γ Glycine Nucleation Revisited: To Stir Or Not To Stir?" In: *CrystEngComm* 21.13 (2019), pp. 2234–2243. DOI: 10.1039/c8ce01829d. URL: https://doi.org/10.1039/c8ce01829d.

[82] Arup Datta and Aslam Hossain. "Solubility Data of Glycine in Water and Justification of Literature Results". In: *Asian J. Chem.* 32.7 (2020), pp. 1525–1533. DOI: 10.14233/ajchem.2020.22626. URL: https://doi.org/10.14233/ajchem.2020.22626.

[83] Lennart A. I. Ramakers et al. "Investigation of Metastable Zones and Induction Times in Glycine Crystallization Across Three Different Antisolvents". In: *Cryst. Growth Des.* 20.8 (2020), pp. 4935–4944. DOI: 10.1021/acs.cgd.9b01493. URL: https://doi.org/10.1021/acs.cgd.9b01493.

[84] Andréanne Bouchard, Gerard W. Hofland, and Geert-Jan Witkamp. "Solubility of Glycine Polymorphs and Recrystallization of β-Glycine". In: *J. Chem. Eng. Data* 52.5 (2007), pp. 1626–1629. DOI: 10.1021/je700014k. URL: https://doi.org/10.1021/je700014k.

[85] Elena N. Tsurko, Roland Neueder, and Werner Kunz. "Water Activity and Osmotic Coefficients in Solutions of Glycine, Glutamic Acid, Histidine and Their Salts At 298.15 K and 310.15 K". In: *J. Solution Chem.* 36.5 (2007), pp. 651–672. DOI: 10.1007/s10953-007-9131-8. URL: https://doi.org/10.1007/s10953-007-9131-8.

[86] X. Xu, S. P. Pinho, and E. A. Macedo. "Activity Coefficient and Solubility of Amino Acids in Water by the Modified Wilson Model". In: *Ind. Eng. Chem. Res.* 43.12 (May 2004), pp. 3200–3204. ISSN: 1520-5045. DOI: 10.1021/ie030860z. URL: http://dx.doi.org/10.1021/ie030860z.

[87] Christoph Held, Luca F. Cameretti, and Gabriele Sadowski. "Measuring and Modeling Activity Coefficients in Aqueous Amino-Acid Solutions". In: *Ind. Eng. Chem. Res.* 50.1 (2011), pp. 131–141. DOI: 10.1021/ie100088c. URL: https://doi.org/10.1021/ie100088c.

[88] K. Renuka Devi, V. Gnanakamatchi, and K. Srinivasan. "Attainment of Unstable β Nucleation of Glycine Through Novel Swift Cooling Crystallization Process". In: *J. Cryst. Growth* 400.nil (2014), pp. 34–42. DOI: 10.1016/j.jcrysgro.2014.04.029. URL: https://doi.org/10.1016/j.jcrysgro.2014.04.029.

[89] Xia Yang, Xiujuan Wang, and Chi Bun Ching. "Solubility of Form α and Form γ of Glycine in Aqueous Solutions". In: *J. Chem. Eng. Data* 53.5 (2008), pp. 1133–1137. DOI: 10.1021/je7006988. URL: https://doi.org/10.1021/je7006988.

[90] K. Igarashi et al. "Control of Polymorphs on the Crystallization of Glycine Using a WWDJ Batch Crystallizer". In: *Eng. Life Sci.* 3.3 (2003), pp. 159–163. DOI: 10.1002/elsc.200390021. URL: https://doi.org/10.1002/elsc.200390021.

[91] Hidetoshi Sakai et al. "Transformation of α-glycine To γ-glycine". In: *J. Cryst. Growth* 116.3-4 (1992), pp. 421–426. DOI: 10.1016/0022-0248(92)90651-x. URL: https://doi.org/10.1016/0022-0248(92)90651-x.

[92] Guangjun Han, Pui Shan Chow, and Reginald B. H. Tan. "Direct Comparison of α- and γ-Glycine Growth Rates in Acidic and Basic Solutions: New Insights Into Glycine Polymorphism". In: *Cryst. Growth Des.* 12.5 (2012), pp. 2213–2220. DOI: 10.1021/cg201321f. URL: https://doi.org/10.1021/cg201321f.

[93] A. Seidell and W. F Linke. *Solubility of Inorganic and Organic Compounds, 3rd ed*. Van Nostrand, 1952.

[94] Hua Sun, Lu Wang, and Baoshu Liu. "Solubility of α-glycine in Water With Additives At a Temperature Range of (293.15-343.15) K: Experimental Data and Results of Thermodynamic Modeling". In: *Fluid Phase Equilib.* 434.nil (2017), pp. 167–175. DOI: 10.1016/j.fluid.2016.12.003. URL: https://doi.org/10.1016/j.fluid.2016.12.003.

[95] Richard Dowling et al. "Acceleration of Crystal Growth Rates: an Unexpected Effect of Tailor-Made Additives". In: *Chem. Commun.* 46.32 (2010), p. 5924. DOI: 10.1039/c0cc00336k. URL: https://doi.org/10.1039/c0cc00336k.

[96] John B. Dalton and Carl L.A. Schmidt. "The Solubilities of Certain Amino Acids in Water, the Densities of Their Solutions At Twenty-Five Degrees, and the Calculated Heats of Solution and Partial Molal Volumes". In: *J. Biol. Chem.* 103.2 (1933), pp. 549–578. DOI: 10.1016/s0021-9258(18)75835-3. URL: https://doi.org/10.1016/s0021-9258(18)75835-3.

[97] Max S. Dunn, Frank J. Ross, and Lee S. Read. "The Solubility of the Amino Acids in Water". In: *J. Biol. Chem.* 103.2 (1933), pp. 579–595. DOI: 10.1016/s0021-9258(18)75836-5. URL: https://doi.org/10.1016/s0021-9258(18)75836-5.

[98] Ziaul Haque Ansari and Zhibao Li. "Solubilities and Modeling of Glycine in Mixed NaCl-MgCl2 Solutions in a Highly Concentrated Region". In: *J. Chem. Eng. Data* 61.10 (2016), pp. 3488–3497. DOI: 10.1021/acs.jced.6b00403. URL: https://doi.org/10.1021/acs.jced.6b00403.

[99] Ziaul Haque Ansari et al. "Modeling of Glycine Solubility in Aqueous HCl-MgCl2 System and Its Application in Phase Transition of Glycine By Changing Media and Supersaturation". In: *J. Cryst. Growth* 467.nil (2017), pp. 116–125. DOI: 10.1016/j.jcrysgro.2017.03.037. URL: https://doi.org/10.1016/j.jcrysgro.2017.03.037.

[100] Yan Zeng, Zhibao Li, and George P. Demopoulos. "Phase Equilibria for the Glycine-Methanol-NH4Cl-H2O System". In: *Ind. Eng. Chem. Res.* 53.43 (2014), pp. 16864–16872. DOI: 10.1021/ie502846m. URL: https://doi.org/10.1021/ie502846m.

[101] Wencheng Gao and Zhibao Li. "Determination and Chemical Modeling of Phase Equilibria for the Glycine-KCl-NaCl-H2O System and Its Application To Produce Crystals With Anticaking Characteristics". In: *Ind. Eng. Chem. Res.* 51.24 (2012), pp. 8315–8325. DOI: 10.1021/ie300787k. URL: https://doi.org/10.1021/ie300787k.

[102] Zheng Cao et al. "Solubility of glycine in binary system of ethanol+water solvent mixtures: Experimental data and thermodynamic modeling". In: *Fluid Phase Equilib.* 360 (Dec. 2013), pp. 156–160. ISSN: 0378-3812. DOI: 10.1016/j.fluid.2013.09.013. URL: http://dx.doi.org/10.1016/j.fluid.2013.09.013.

[103] J. W Mullin. *Crystallization*. Oxford Boston: Butterworth-Heinemann, 2001. ISBN: 9780750648332.

[104] Leping Dang et al. "The Effect of Temperature and Solvent Composition on Transformation of β- To α-Glycine As Monitored in Situ By FBRM and PVM". In: *Org. Process Res. Dev.* 13.6 (2009), pp. 1301–1306. DOI: 10.1021/op900134w. URL: https://doi.org/10.1021/op900134w.

[105] G. L. Perlovich, L. K. Hansen, and A. Bauer-Brandl. "The polymorphism of glycine: Thermochemical and structural aspects". In: *J. Therm. Anal. Calorim.* 66.3 (2001), pp. 699–715. ISSN: 1418-2874. DOI: 10.1023/a:1013179702730. URL: http://dx.doi.org/10.1023/A:1013179702730.

[106] Emmerich Wilhelm. "Mitigating Complexity: Cohesion Parameters and Related Topics. I: the Hildebrand Solubility Parameter". In: *J. Solution Chem.* 47.10 (2018), pp. 1626–1709. DOI: 10.1007/s10953-018-0821-1. URL: https://doi.org/10.1007/s10953-018-0821-1.

[107] Hidetoshi Kuramochi et al. "Measurements of Vapor Pressures of Aqueous Amino Acid Solutions and Determination of Activity Coefficients of Amino Acids". In: *J. Chem. Eng. Data* 42.3 (May 1997), pp. 470–474. ISSN: 1520-5134. DOI: 10.1021/je960113r. URL: http://dx.doi.org/10.1021/je960113r.

[108] V. E. Bower and R. A. Robinson. "Thermodynamics of the Ternary System: Water-Glycine-Potassium Chloride At 25°C From Vapor Pressure Measurements". In: *J. Res. Natl. Bur. Stand., A Phys. Chem.* 69A.2 (1965), p. 131. DOI: 10.6028/jres.069a.015. URL: https://doi.org/10.6028/jres.069a.015.

[109] H. David Ellerton et al. "Activity, Density, and Relative Viscosity Data for Several Amino Acids, Lactamide, and Raffinose in Aqueous Solution At 25°C". In: *J. Phys. Chem.* 68.2 (1964), pp. 398–402. DOI: 10.1021/j100784a034. URL: https://doi.org/10.1021/j100784a034.

[110] Deepti N. Kurhe et al. "Studies of Enthalpy-Entropy Compensation, Partial Entropies, and Kirkwood-Buff Integrals for Aqueous Solutions of Glycine, L-Leucine, and Glycylglycine At 298.15 K". In: *J. Phys. Chem. B* 113.52 (2009), pp. 16612–16621. DOI: 10.1021/jp9078585. URL: https://doi.org/10.1021/jp9078585.

[111] Elizabeth R.B. Smith and Paul K. Smith. "The Activity of Glycine in Aqueous Solution At Twenty-Five Degrees". In: *J. Biol. Chem.* 117.1 (1937), pp. 209–216. DOI: 10.1016/s0021-9258(18)74602-4. URL: https://doi.org/10.1016/s0021-9258(18)74602-4.

[112] Terence H. Lilley and R. P. Scott. "Aqueous Solutions Containing Amino-Acids and Peptides. Part 2.-gibbs Function and Enthalpy Behaviour of the Systems Urea + Glycine, Urea +α-alanine, Urea +α-aminobutyric Acid and Urea + Glycylglycine At 298.15 K". In: *J. Chem. Soc., Faraday Trans. 1 F* 72.0 (1976), p. 184. DOI: 10.1039/f19767200184. URL: https://doi.org/10.1039/f19767200184.

[113] Alireza Salabat, Shakiba Neshat, and Alireza Fazlali. "Activity Coefficients of Glycine, D-Alanine and L-Valine in Aqueous Solutions Containing MgSO4 At 298.15 K; Experimental Determination and Correlation". In: *Fluid Phase Equilib.* 314.nil (2012), pp. 198–202. DOI: 10.1016/j.fluid.2011.10.015. URL: https://doi.org/10.1016/j.fluid.2011.10.015.

[114] Simão P. Pinho. "Water Activity in Aqueous Amino Acid Solutions, With and Without KCl, At 298.15 K". In: *J. Chem. Eng. Data* 53.1 (2007), pp. 180–184. DOI: 10.1021/je700466s. URL: https://doi.org/10.1021/je700466s.

[115] L. Ninni and A.J.A. Meirelles. "Water Activity, Ph and Density of Aqueous Amino Acids Solutions". In: *Biotechnol. Progr.* 17.4 (2001), pp. 703–711. DOI: 10.1021/bp0100427. URL: https://doi.org/10.1021/bp0100427.

[116] Rahmat Sadeghi and Afsaneh Gholamireza. "Thermodynamics of the Ternary Systems: (water + Glycine, L-Alanine and L-Serine + Di-Ammonium Hydrogen citrate) From Volumetric, Compressibility, and (vapour + liquid) Equilibria Measurements". In: *J. Chem. Thermodyn.* 43.2 (2011), pp. 200–215. DOI: 10.1016/j.jct.2010.08.021. URL: https://doi.org/10.1016/j.jct.2010.08.021.

[117] Julian M. Sturtevant. "The Heats of Dilution of Aqueous Solutions of Glycine At 25°C". In: *J. Am. Chem. Soc.* 62.7 (1940), pp. 1879–1879. DOI: 10.1021/ja01864a507. URL: https://doi.org/10.1021/ja01864a507.

[118] R.S. Humphrey et al. "The Partial Molar Enthalpies in Aqueous Solution of Some Amino Acids With Polar and Non-Polar Side Chains". In: *J. Chem. Thermodyn.* 12.6 (1980), pp. 595–603. DOI: 10.1016/0021-9614(80)90191-3. URL: https://doi.org/10.1016/0021-9614(80)90191-3.

[119] Frank T. Gucker, Hugh B. Pickard, and William L. Ford. "The Heats of Dilution of Aqueous Solutions of Glycine and Glycolamide, and Other Thermodynamic Properties of Glycine at 25°C". In: *J. Am. Chem. Soc.* 62.10 (Oct. 1940), pp. 2698–2704. ISSN: 1520-5126. DOI: 10.1021/ja01867a027. URL: http://dx.doi.org/10.1021/ja01867a027.

[120] Xu Wang et al. "Enthalpies of Dilution of Glycine, L-Alanine and L-Serine in Aqueous Potassium Chloride Solutions". In: *Thermochim. Acta* 425.1-2 (2005), pp. 31–37. DOI: 10.1016/j.tca.2004.05.021. URL: https://doi.org/10.1016/j.tca.2004.05.021.

[121] W. E. Wallace, William F. Offutt, and A. L. Robinson. "The Heats of Dilution of Aqueous Solutions of Glycine At 25°C". In: *J. Am. Chem. Soc.* 65.3 (1943), pp. 347–350. DOI: 10.1021/ja01243a014. URL: https://doi.org/10.1021/ja01243a014.

[122] B. Pałecz. "Enthalpies os Solution and Dilution of Some L-α-Amino Acids in Water at 298.15 K". In: *J. Therm. Anal. Calorim.* 54.1 (1998), pp. 257–263. DOI: 10.1023/a:1010158011499. URL: https://doi.org/10.1023/a:1010158011499.

[123] Shuqin Li et al. "Enthalpic Interaction of Glycine in Aqueous Glucose and Sucrose Solutions At 298.15 K". In: *Thermochim. Acta* 342.1-2 (1999), pp. 1–6. DOI: 10.1016/s0040-6031(99)00111-2. URL: https://doi.org/10.1016/s0040-6031(99)00111-2.

[124] Xiaoling Ren, Yaming Ni, and Ruisen Lin. "Enthalpies of Dilution of Glycine, L-Serine and L-Valine in Mixtures of Water and N,n-Dimethylformamide At 298.15 K". In: *Thermochim. Acta* 348.1-2 (2000), pp. 19–24. DOI: 10.1016/s0040-6031(99)00358-5. URL: https://doi.org/10.1016/s0040-6031(99)00358-5.

[125] Frank T. Gucker and Theodore W. Allen. "The Densities and Specific Heats of Aqueous Solutions of dl-α-Alanine, β-Alanine and Lactamide1,2". In: *J. Am. Chem. Soc.* 64.2 (Feb. 1942), pp. 191–199. ISSN: 1520-5126. DOI: 10.1021/ja01254a001. URL: http://dx.doi.org/10.1021/ja01254a001.

[126] Colin J. Downes, Andrew W. Hakin, and Gavin R. Hedwig. "The partial molar heat capacities of glycine and glycylglycine in aqueous solution at elevated temperatures and at p=10.0 MPa". In: *J. Chem. Thermodyn.* 33.8 (Aug. 2001), pp. 873–890. ISSN: 0021-9614. DOI: 10.1006/jcht.2000.0796. URL: http://dx.doi.org/10.1006/jcht.2000.0796.

[127] C.H. Spink and I. Wadsö. "Thermochemistry of solutions of biochemical model compounds. 4. The partial molar heat capacities of some amino acids in aqueous solution". In: *J. Chem. Thermodyn.* 7.6 (June 1975), pp. 561–572. ISSN: 0021-9614. DOI: 10.1016/0021-9614(75)90190-1. URL: http://dx.doi.org/10.1016/0021-9614(75)90190-1.

[128] Giuseppa DiPaola and Bernard Belleau. "Apparent Molal Heat Capacities and Volumes of Amino Acids in Aqueous Polyol Solutions". In: *Can. J. Chem.* 56.13 (1978), pp. 1827–1831. DOI: 10.1139/v78-296. URL: https://doi.org/10.1139/v78-296.

[129] V. A. Drebushchak et al. "Low-Temperature Heat Capacity of β-glycine and a Phase Transition At 252 K". In: *J. Therm. Anal. Calorim.* 79.1 (2005), pp. 65–70. DOI: 10.1007/s10973-004-0563-8. URL: https://doi.org/10.1007/s10973-004-0563-8.

[130] V. A. Drebushchak, A. G. Ogienko, and Elena V. Boldyreva. "Polymorphic Effects At the Eutectic Melting in the H2O-Glycine System". In: *J. Therm. Anal. Calorim.* 111.3 (2012), pp. 2187–2194. DOI: 10.1007/s10973-012-2761-0. URL: https://doi.org/10.1007/s10973-012-2761-0.

[131] G. A. Anslow, M. L. Foster, and C. Klinger. "The Absorption Spectra of Glycine Solutions and Their Interpretation". In: *J. Biol. Chem.* 103 (1933), pp. 81–92.

[132] W. C. M. Lewis. "The Crystallization, Denaturation and Flocculation of Proteins With Special Reference To Albumin and Hemoglobin; Together With an Appendix on the Physicochemical Behavior of Glycine." In: *Chem. Rev.* 8.1 (1931), pp. 81–165. DOI: 10.1021/cr60029a002. URL: https://doi.org/10.1021/cr60029a002.

[133] Jun Huang, Thomas C. Stringfellow, and Lian Yu. "Glycine Exists Mainly As Monomers, Not Dimers, in Supersaturated Aqueous Solutions: Implications for Understanding Its Crystallization and Polymorphism". In: *J. Am. Chem. Soc.* 130.42 (2008), pp. 13973–13980. DOI: 10.1021/ja804836d. URL: https://doi.org/10.1021/ja804836d.

[134] M. Shimoyamada et al. "Freezing and Eutectic Points of an Aqueous Amino Acid Solution Containing Ethanol, and the Effect of Ethanol Addition on the Freeze Concentration Process". In: *Biosci Biotechnol Biochem* 58.5 (1994), pp. 836–838. DOI: 10.1271/bbb.58.836. URL: https://doi.org/10.1271/bbb.58.836.

[135] George Scatchard and S. S. Prentiss. "Freezing Points of Aqueous Solutions. VII. Ethyl Alcohol, Glycine and Their Mixtures". In: *J. Am. Chem. Soc.* 56.7 (1934), pp. 1486–1492. DOI: 10.1021/ja01322a013. URL: https://doi.org/10.1021/ja01322a013.

[136] V. A. Drebushchak et al. "Low-temperature heat capacity of α and γ polymorphs of glycine". In: *J. Therm. Anal. Calorim.* 74.1 (2003), pp. 109–120. DOI: 10.1023/a:1026377703260. URL: https://doi.org/10.1023/a:1026377703260.

[137] K. Srinivasan. "Crystal Growth of α and γ Glycine Polymorphs and Their Polymorphic Phase Transformations". In: *J. Cryst. Growth* 311.1 (2008), pp. 156–162. DOI: 10.1016/j.jcrysgro.2008.10.084. URL: https://doi.org/10.1016/j.jcrysgro.2008.10.084.

[138] Y. Iitaka. "The crystal structure of γ-glycine". In: *Acta Crystallogr.* 14.1 (Jan. 1961), pp. 1–10. ISSN: 0365-110X. DOI: 10.1107/s0365110x61000012. URL: http://dx.doi.org/10.1107/S0365110X61000012.

[139] V. M. Kozhin. "Thermal expansion tensors of α-, β-, and γ-modifications of glycine". In: *Kristallografiya* 23.6 (1978), pp. 1211–15.

[140] T. Balakrishnan, R. Ramesh Babu, and K. Ramamurthi. "Growth, Structural, Optical and Thermal Properties of γ-glycine Crystal". In: *Spectrochim. Acta, Part A* 69.4 (2008), pp. 1114–1118. DOI: 10.1016/j.saa.2007.06.025. URL: https://doi.org/10.1016/j.saa.2007.06.025.

[141] M.Narayan Bhat and S.M Dharmaprakash. "Growth of Nonlinear Optical γ-glycine Crystals". In: *J. Cryst. Growth* 236.1-3 (2002), pp. 376–380. DOI: 10.1016/s0022-0248(01)02094-2. URL: https://doi.org/10.1016/s0022-0248(01)02094-2.

[142] E. V. Boldyreva et al. "Polymorphism of glycine, Part II". In: *J. Therm. Anal. Calorim.* 73.2 (2003), pp. 419–428. DOI: 10.1023/a:1025457524874. URL: https://doi.org/10.1023/a:1025457524874.

[143] A. T. Williamson. "The exact calculation of heats of solution from solubility data". In: *Transactions of the Faraday Society* 40 (1944), p. 421. DOI: 10.1039/tf9444000421. URL: https://doi.org/10.1039/tf9444000421.

[144] Simon Black and Frans Muller. "On the Effect of Temperature on Aqueous Solubility of Organic Solids". en. In: *Organic Process Research & Development* 14.3 (May 2010), pp. 661–665. ISSN: 1083-6160, 1520-586X. DOI: 10.1021/op100006y. URL: https://pubs.acs.org/doi/10.1021/op100006y (visited on 01/26/2022).

[145] M Starzak and M Mathlouthi. "Temperature dependence of water activity in aqueous solutions of sucrose". In: *Food Chemistry* 96.3 (June 2006), pp. 346–370. DOI: 10.1016/j.foodchem.2005.02.052. URL: https://doi.org/10.1016/j.foodchem.2005.02.052.

[146] G. Grube and M. Nussbaum. "Phasentheoretische Untersuchungen über die Entzuckerung der Melasse.: I. Mitteilung: das ternäre System Strontiumoxyd-Saccharose-Wasser". In: *Zeitschrift fur Elektrochemie und angewandte phyysikalische Chemie* 34 (1928), pp. 91–98.

[147] J. R. Vernon. *The velocity of crystallization of sucrose*. University of London, 1939.

[148] Frank E. Young and Francis T. Jones. "Sucrose Hydrates. The Sucrose–Water Phase Diagram." In: *The Journal of Physical and Colloid Chemistry* 53.9 (Sept. 1949), pp. 1334–1350. DOI: 10.1021/j150474a004. URL: https://doi.org/10.1021/j150474a004.

[149] W. S. Wise and E. B. Nicholson. "The solubilities and heats of crystallisation of sucrose and methyl α-D-glucoside in aqueous solution". In: *J. Chem. Soc.* 0.0 (1955), pp. 2714–2716. DOI: 10.1039/jr9550002714. URL: https://doi.org/10.1039/jr9550002714.

[150] H. Rother. *Über den Dampfdruck von Saccharose- and Glucose-Lösungen. Ein Beitrag zur Thermodynamik von Nichtelektrolytlösungen*. Hochschule, Braunschweig, 1960.

[151] D. F. Charles. "Solubility of pure sucrose in water". In: *Int. Sugar J* 62 (1960), p. 126.

[152] L. A. Pinck and Mary A. Kelly. "The Solubility of Urea in Water." In: *Journal of the American Chemical Society* 47.8 (Aug. 1925), pp. 2170–2172. DOI: 10.1021/ja01685a502. URL: https://doi.org/10.1021/ja01685a502.

[153] Louis Shnidman and A. A. Sunier. "The Solubility of Urea in Water". In: *The Journal of Physical Chemistry* 36.4 (Apr. 1932), pp. 1232–1240. DOI: 10.1021/j150334a013. URL: https://doi.org/10.1021/j150334a013.

[154] Lawrence H. Dalman. "Ternary Systems of Urea and Acids. IV. Urea, Citric Acid and Water. V. Urea, Acetic Acid and Water. VI. Urea, Tartaric Acid and Water". In: *Journal of the American Chemical Society* 59.5 (May 1937), pp. 775–779. DOI: 10.1021/ja01284a001. URL: https://doi.org/10.1021/ja01284a001.

[155] Atherton Seidell and William F. Linke. *Solubilities of inorganic and organic compounds*. Vol. 4. 1952.

[156] Harper Kakinuma. "Communication to the Editor: The Solubility of Urea in Water." In: *The Journal of Physical Chemistry* 45.6 (June 1941), pp. 1045–1046. DOI: 10.1021/j150411a014. URL: https://doi.org/10.1021/j150411a014.

[157] Fu-Ming Lee and Leslie E. Lahti. "Solubility of urea in water-alcohol mixtures". In: *Journal of Chemical & Engineering Data* 17.3 (July 1972), pp. 304–306. DOI: 10.1021/je60054a020. URL: https://doi.org/10.1021/je60054a020.

[158] Jing Chen et al. "Determination and Correlation of Solubility of Decahydropyrazino[2, 3-b]pyrazine in Methanol, Ethanol, and 2-Propanol". In: *Industrial & Engineering Chemistry Research* 50.20 (Sept. 2011), pp. 11755–11762. DOI: 10.1021/ie2012714. URL: https://doi.org/10.1021/ie2012714.

[159] John M Prausnitz and etc. *Molecular thermodynamics of fluid-phase equilibria*. en. 2nd ed. Prentice-Hall international series in the physical and chemical engineering sciences. Philadelphia, PA: Prentice Hall, Oct. 1985.

[160] J. A. Nelder and R. Mead. "A Simplex Method for Function Minimization". In: *The Computer Journal* 7.4 (Jan. 1965), pp. 308–313. DOI: 10.1093/comjnl/7.4.308. URL: https://doi.org/10.1093/comjnl/7.4.308.

[161] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[162] Stephen A Cooke, Svava Ósk Jónsdóttir, and Peter Westh. "The vapour pressure of water as a function of solute concentration above aqueous solutions of fructose, sucrose, raffinose, erythritol, xylitol, and sorbitol". In: *The Journal of Chemical Thermodynamics* 34.10 (Oct. 2002), pp. 1545–1555. DOI: 10.1016/s0021-9614(02)00172-6. URL: https://doi.org/10.1016/s0021-9614(02)00172-6.

[163] Jose F. Comesana, Antonio Correa, and Alberto M. Sereno. "Water activity at 35 oC in 'sugar ' + water and 'sugar ' + sodium chloride + water systems". In: *International Journal of Food Science and Technology* 36.6 (Aug. 2001), pp. 655–661. DOI: 10.1046/j.1365-2621.2001.00501.x. URL: https://doi.org/10.1046/j.1365-2621.2001.00501.x.

[164] H. I. Downes and E. P. Perman. "Vapour pressure and heat of dilution of aqueous solutions, part IV. An improved method of measuring vapour pressure by air bubbling". In: *Transactions of the Faraday Society* 23 (1927), 95b. DOI: 10.1039/tf927230095b. URL: https://doi.org/10.1039/tf927230095b.

[165] R Fricke and L Havestadt. "Work of dilution and heat of dilution in the sphere of concentrated solutions". In: *Z. Elektroehem. angew. phys. Chem.* 33 (1927), pp. 441–455.

[166] D. W. Grover. "The keeping properties of confectionery as influenced by its water vapour pressure". In: *Journal of the Society of Chemical Industry* 66.7 (July 1947), pp. 201–205. DOI: 10.1002/jctb.5000660701. URL: https://doi.org/10.1002/jctb.5000660701.

[167] Bahman Jamehbozorg and Rahmat Sadeghi. "Vapor Pressure Osmometry Studies of Aqueous Ionic Liquid–Carbohydrate Systems". In: *Journal of Chemical & Engineering Data* 63.2 (Jan. 2018), pp. 331–340. DOI: 10.1021/acs.jced.7b00719. URL: https://doi.org/10.1021/acs.jced.7b00719.

[168] F. C. Krauskopf. *The Vapor Pressure of Water and Aqueous Solutions of Sodium Chloride, Potassium Chloride and Sugar*. Springer Handbooks. University of Wisconsin–Madison, 1910.

[169] R. W. Money and R. Born. "Equilibrium humidity of sugar solutions". In: *Journal of the Science of Food and Agriculture* 2.4 (Apr. 1951), pp. 180–185. DOI: 10.1002/jsfa.2740020408. URL: https://doi.org/10.1002/jsfa.2740020408.

[170] Robert A. Robinson, Paul K. Smith, and Elizabeth R. B. Smith. "The osmotic coefficients of some organic compounds in relation to their chemical constitution". In: *Transactions of the Faraday Society* 38 (1942), p. 63. DOI: 10.1039/tf9423800063. URL: https://doi.org/10.1039/tf9423800063.

[171] George Scatchard, W. J. Hamer, and S. E. Wood. "Isotonic Solutions. I. The Chemical Potential of Water in Aqueous Solutions of Sodium Chloride, Potassium Chloride, Sulfuric Acid, Sucrose, Urea and Glycerol at 25C". In: *Journal of the American Chemical Society* 60.12 (Dec. 1938), pp. 3061–3070. DOI: 10.1021/ja01279a066. URL: https://doi.org/10.1021/ja01279a066.

[172] D. A. Sinclair. "A Simple Method for Accurate Determinations of Vapor Pressures of Solutions." In: *The Journal of Physical Chemistry* 37.4 (Apr. 1933), pp. 495–504. DOI: 10.1021/j150346a011. URL: https://doi.org/10.1021/j150346a011.

[173] Carmen E. Velezmoro et al. "Prediction of Water Activity in Sugar Solutions Using Models of Group Contribution and Equation of State." In: *JOURNAL OF CHEMICAL ENGINEERING OF JAPAN* 33.4 (2000), pp. 645–653. DOI: 10.1252/jcej.33.645. URL: https://doi.org/10.1252/jcej.33.645.

[174] Mohammed Taghi Zafarani-Moattar, Hemayat Shekaari, and Elnaz Mazaher Haji Agha. "Vapor – Liquid equilibria study of the ternary systems containing sucrose in aqueous solutions of ionic liquids, 1-butyl-3-methyl imidazolium bromide and 1-hexyl-3-methyl imidazolium bromide at 298.15 K and atmospheric pressure". In: *Fluid Phase Equilibria* 429 (Dec. 2016), pp. 45–54. DOI: 10.1016/j.fluid.2016.08.025. URL: https://doi.org/10.1016/j.fluid.2016.08.025.

[175] Robert H. Wood and Laurence H. Hiltzik. "Enthalpies of dilution of aqueous solutions of formamide, acetamide, propionamide, andN, N-dimethylformamide". In: *Journal of Solution Chemistry* 9.1 (Jan. 1980), pp. 45–57. DOI: 10.1007/bf00650136. URL: https://doi.org/10.1007/bf00650136.

[176] Freda M. Hunter. "Latent heat of dilution of cane sugar solutions". In: *Transactions of the Faraday Society* 22 (1926), p. 194. DOI: 10.1039/tf9262200194. URL: https://doi.org/10.1039/tf9262200194.

[177] K. Bennewitz and L Kratz. In: *Phys. Z* 37 (1936), pp. 496–511.

[178] J. D'ans and H. Tollert. "Die Bestimmung der spezifischen Warmen konzentrierter wassriger Salzlosungen". In: *Z. Elektrochem.* 43.2 (1937), pp. 81–91.

[179] C. Frenzel, R. Burian, and O. Haas. "Verdunnungswarmen und osmotische Drucke von Nichtelektrolytlosungen". In: *Z. Elektrochem.* 41 (7 1935), pp. 419–429.

[180] Frank T. Gucker and Fred D. Ayres. "The Specific Heats of Aqueous Sucrose Solutions at 20 and 25° and the Apparent Molal Heat Capacity of Non-electrolytes". In: *Journal of the American Chemical Society* 59.3 (Mar. 1937), pp. 447–452. DOI: 10.1021/ja01282a006. URL: https://doi.org/10.1021/ja01282a006.

[181] H. Hausrath. "A differential method for determining small freezing point depressions". In: *Ann. Phys. (Leipzig)* 314.11 (1902), pp. 522–554.

[182] A. D. C. Rivett. "Influence of neutral salts on the freezing point of mixtures in aqueous solution". In: *Z. Phys. chem. Leipzig* 80 (1912), pp. 537–563.

[183] K. Kiyosawa. "The volumes of hydrated glucose, sucrose and raffinose molecules, and the osmotic pressures of these aqueous saccharide solutions as measured by the freezing-point-depression method". In: *Bull. chem. Soc. Jpn* 61 (3 1988), pp. 633–642.

[184] Massimo Migliori et al. "Viscosity of Multicomponent Solutions of Simple and Complex Sugars in Water". In: *Journal of Chemical & Engineering Data* 52.4 (June 2007), pp. 1347–1353. DOI: 10.1021/je700062x. URL: https://doi.org/10.1021/je700062x.

[185] V. E. Bower and R. A. Robinson. "THE THERMODYNAMICS OF THE TERNARY SYSTEM: UREA-SODIUM CHLORIDE-WATER AT 25°". In: *The Journal of Physical Chemistry* 67.7 (July 1963), pp. 1524–1527. DOI: 10.1021/j100801a030. URL: https://doi.org/10.1021/j100801a030.

[186] H. David Ellerton and Peter J. Dunlop. "Activity Coefficients for the Systems Water-Urea and Water-Urea-Sucrose at 25° from Isopiestic Measurements". In: *The Journal of Physical Chemistry* 70.6 (June 1966), pp. 1831–1837. DOI: 10.1021/j100878a023. URL: https://doi.org/10.1021/j100878a023.

[187] CW Childs and RF Platford. "Excess free energies of mixing at temperatures below 25°. Isopiestic measurements on the systems H2O-NaCl-Na2SO4 and H2O-NaCl-MgSO4". In: *Australian Journal of Chemistry* 24.12 (1971), p. 2487. DOI: 10.1071/ch9712487. URL: https://doi.org/10.1071/ch9712487.

[188] Deepti N. Kurhe et al. "Thermodynamic Studies of Amino Acid–Denaturant Interactions in Aqueous Solutions at 298.15 K". In: *Journal of Solution Chemistry* 40.9 (Sept. 2011), pp. 1596–1617. DOI: 10.1007/s10953-011-9737-8. URL: https://doi.org/10.1007/s10953-011-9737-8.

[189] Sopan K. Kushare et al. "Studies of the Effect of Urea on PEG-4000 Polymer–Water Interactions at 298.15 K". In: *Journal of Chemical Engineering Data* 64.6 (Apr. 2019), pp. 2341–2349. DOI: 10.1021/acs.jced.8b01121. URL: https://doi.org/10.1021/acs.jced.8b01121.

[190] Frank T. Gucker and Hugh B. Pickard. "The Heats of Dilution, Heat Capacities, and Activities of Urea in Aqueous Solutions from the Freezing Points to 40°". In: *Journal of the American Chemical Society* 62.6 (June 1940), pp. 1464–1472. DOI: 10.1021/ja01863a038. URL: https://doi.org/10.1021/ja01863a038.

[191] D. Hamilton and R. H. Stokes. "Enthalpies of dilution of urea solutions in six polar solvents at several temperatures". In: *Journal of Solution Chemistry* 1.3 (Sept. 1972), pp. 223–235. DOI: 10.1007/bf00645103. URL: https://doi.org/10.1007/bf00645103.

[192] J Malec and M Sarnowski. "Integral Heat of Solution of Urea in Water and in Calcium & Chloride Aqueous Solutions at 25 +- 0.75 C". In: *Pol. J. Chem* 45.9 (1971), pp. 1369–1375.

[193] Goolam M Musbally, Gerald Perron, and Jacques E Desnoyers. "The effect of temperature and urea on the apparent molal volumes and heat capacities of n-nonyltrimethylammonium bromide in water". In: *Journal of Colloid and Interface Science* 54.1 (Jan. 1976), pp. 80–93. DOI: 10.1016/0021-9797(76)90287-3. URL: https://doi.org/10.1016/0021-9797(76)90287-3.

[194] E. P. Egan and B. B. Luff. "Heat of Solution, Heat Capacity, and Density of Aqueous Urea Solutions at 25° C." In: *Journal of Chemical & Engineering Data* 11.2 (Apr. 1966), pp. 192–194. DOI: 10.1021/je60029a020. URL: https://doi.org/10.1021/je60029a020.

[195] Andrew W. Hakin, Colin L. Beswick, and Michelle M. Duke. "Thermochemical and volumetric properties of aqueous urea systems. Heat capacities and volumes of transfer from water to urea—water mixtures for some 1 : 1 electrolytes at 298.15 K". In: *J. Chem. Soc., Faraday Trans.* 92.2 (1996), pp. 207–213. DOI: 10.1039/ft9969200207. URL: https://doi.org/10.1039/ft9969200207.

[196] b.r. brown et al. "apparent molar volumes and apparent molar heat capacities of aqueous urea, 1, 1-dimethylurea, and n, n′-dimethylurea at temperatures from (278.15 to 348.15)k and at the pressure 0.35mpa". In: *the journal of chemical thermodynamics* 38.8 (Aug. 2006), pp. 1025–1035. DOI: 10.1016/j.jct.2005.10.017. URL: https://doi.org/10.1016/j.jct.2005.10.017.

[197] H. M. Chadwell and F. W. Politi. "The Freezing Points of Concentrated Aqueous Solutions of Urea, Urethan, and Acetamide". In: *Journal of the American Chemical Society* 60.6 (June 1938), pp. 1291–1293. DOI: 10.1021/ja01273a004. URL: https://doi.org/10.1021/ja01273a004.

[198] Hans Lukas, Suzana G Fries, and Bo Sundman. *Computational thermodynamics*. Cambridge, England: Cambridge University Press, July 2007.

[199] P.J. Spencer. "A brief history of CALPHAD". In: *Calphad* 32.1 (Mar. 2008), pp. 1–8. DOI: 10.1016/j.calphad.2007.10.001. URL: https://doi.org/10.1016/j.calphad.2007.10.001.

[200] P.J. Spencer. "The origins, growth and current industrial impact of Calphad". In: *Calphad* 79 (Dec. 2022), p. 102489. DOI: 10.1016/j.calphad.2022.102489. URL: https://doi.org/10.1016/j.calphad.2022.102489.

[201] Ales Kroupa. "Modelling of phase diagrams and thermodynamic properties using Calphad method – Development of thermodynamic databases". In: *Computational Materials Science* 66 (Jan. 2013), pp. 3–13. DOI: 10.1016/j.commatsci.2012.02.003. URL: https://doi.org/10.1016/j.commatsci.2012.02.003.

[202] N (nigel) Saunders. *CALPHAD (calculation of phase diagrams): A comprehensive guide: Volume 1*. en. Ed. by N Saunders and A P Miodownik. Pergamon materials series. London, England: Pergamon Press, June 1998.

[203] Mats Hillert. *Phase equilibria, phase diagrams and phase transformations*. 2nd ed. Cambridge, England: Cambridge University Press, June 2012.

[204] Brandon Bocklund et al. "ESPEI for efficient thermodynamic database development, modification, and uncertainty quantification: application to Cu–Mg". In: *MRS Communications* 9.2 (June 2019), pp. 618–627. DOI: 10.1557/mrc.2019.59. URL: https://doi.org/10.1557/mrc.2019.59.

[205] George Kaptay. "A new equation for the temperature dependence of the excess Gibbs energy of solution phases". In: *Calphad* 28.2 (June 2004), pp. 115–124. DOI: 10.1016/j.calphad.2004.08.005. URL: https://doi.org/10.1016/j.calphad.2004.08.005.

[206] George Kaptay. "On the abilities and limitations of the linear, exponential and combined models to describe the temperature dependence of the excess Gibbs energy of solutions". In: *Calphad* 44 (Mar. 2014), pp. 81–94. DOI: 10.1016/j.calphad.2013.08.007. URL: https://doi.org/10.1016/j.calphad.2013.08.007.

[207] George Kaptay. "The exponential excess Gibbs energy model revisited". In: *Calphad* 56 (Mar. 2017), pp. 169–184. DOI: 10.1016/j.calphad.2017.01.002. URL: https://doi.org/10.1016/j.calphad.2017.01.002.

[208] Song-Mao Liang, Peisheng Wang, and Rainer Schmid-Fetzer. "Inherently consistent temperature function for interaction parameters demonstrated for the Mg–Si assessment". In: *Calphad* 54 (Sept. 2016), pp. 82–96. DOI: 10.1016/j.calphad.2016.06.003. URL: https://doi.org/10.1016/j.calphad.2016.06.003.

[209] Jing Li et al. "Implementation of the UNIQUAC model in the OpenCalphad software". In: *Fluid Phase Equilibria* 507 (Mar. 2020), p. 112398. DOI: 10.1016/j.fluid.2019.112398. URL: https://doi.org/10.1016/j.fluid.2019.112398.

[210] Bo Sundman et al. "OpenCalphad - a free thermodynamic software". In: *Integrating Materials and Manufacturing Innovation* 4.1 (Jan. 2015), pp. 1–15. DOI: 10.1186/s40192-014-0029-1. URL: https://doi.org/10.1186/s40192-014-0029-1.

[211] M. DuBois and P. Berge. "Experimental Study of Rayleigh Scattering Related to Concentration Fluctuations in Binary Solutions: Evidence of a Departure from Ideality". In: *Phys. Rev. Lett.* 26 (3 Jan. 1971), pp. 121–124. DOI: 10.1103/PhysRevLett.26.121. URL: https://link.aps.org/doi/10.1103/PhysRevLett.26.121.

[212] A.T. Dinsdale. "SGTE data for pure elements". In: *Calphad* 15.4 (Oct. 1991), pp. 317–425. DOI: 10.1016/0364-5916(91)90030-n. URL: https://doi.org/10.1016/0364-5916(91)90030-n.

[213] Esther Forte et al. "Multi-criteria optimization for parametrizing excess Gibbs energy models". In: *Fluid Phase Equilibria* 522 (Nov. 2020), p. 112676. DOI: 10.1016/j.fluid.2020.112676. URL: https://doi.org/10.1016/j.fluid.2020.112676.

[214] Richard D. Neidinger. "Introduction to Automatic Differentiation and MATLAB Object-Oriented Programming". In: *SIAM Review* 52.3 (Jan. 2010), pp. 545–563. DOI: 10.1137/080743627. URL: https://doi.org/10.1137/080743627.

[215] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

[216] The pandas development team. *pandas-dev/pandas: Pandas*. Version latest. Feb. 2020. DOI: 10.5281/zenodo.3509134. URL: https://doi.org/10.5281/zenodo.3509134.

[217] Wes McKinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[218] H. Claassen. "Bestimmung der Siedepunkte reiner und unreiner Lösungen." In: *Zeitschrift des Vereins für die Rübenzuckerindustrie des Deutschen Reiches* 54 (1904), pp. 1159–1169.

[219] Ewald Plake. "Ein verbessertes Differentialebullioskop". In: *Zeitschrift für Physikalische Chemie* 172A.1 (Jan. 1935), pp. 105–112. DOI: 10.1515/zpch-1935-17209. URL: https://doi.org/10.1515/zpch-1935-17209.

[220] Alfred Holven. "Sucrose Solutions: Influence of Pressure on Boiling Point Elevation". In: *Industrial & Engineering Chemistry* 28.4 (Apr. 1936), pp. 452–455. DOI: 10.1021/ie50316a021. URL: https://doi.org/10.1021/ie50316a021.

[221] C. J. Tressler, W. I. Zimmerman, and C. O. Willits. "Boiling-point Elevation of Sucrose Solutions." In: *The Journal of Physical Chemistry* 45.8 (Aug. 1941), pp. 1242–1245. DOI: 10.1021/j150413a010. URL: https://doi.org/10.1021/j150413a010.

[222] VI Tuzhilkin and IN Kaganov. "Boiling point temperatures of concentrated sugar solutions". In: *Sakharnaya Promyshlennost* 43.1 (1969), pp. 17–20.

[223] S. Abderafi and T. Bounahmidi. "Measurement and modeling of atmospheric pressure vapor-liquid equilibrium data for binary, ternary and quaternary mixtures of sucrose, glucose, fructose and water components". In: *Fluid Phase Equilibria* 93 (Feb. 1994), pp. 337–351. DOI: 10.1016/0378-3812(94)87017-9. URL: https://doi.org/10.1016/0378-3812(94)87017-9.

[224] Y. Abed et al. "Measurement of liquid-solid phase equilibrium in ternary systems of water-sucrose-glucose and water-sucrose fructose, and predictions with UNIFAC". In: *Fluid Phase Equilibria* 73.1-2 (Apr. 1992), pp. 175–184. DOI: 10.1016/0378-3812(92)85047-c. URL: https://doi.org/10.1016/0378-3812(92)85047-c.

[225] S. Abderafi and T. Bounahmidi. "Measurement and modeling of atmospheric pressure vapor-liquid equilibrium data for binary, ternary and quaternary mixtures of sucrose, glucose, fructose and water components". In: *Fluid Phase Equilibria* 93 (Feb. 1994), pp. 337–351. DOI: 10.1016/0378-3812(94)87017-9. URL: https://doi.org/10.1016/0378-3812(94)87017-9.

[226] Marianne Catté et al. "Excess properties and solid-liquid equilibria for aqueous solutions of sugars using a UNIQUAC model". In: *Fluid Phase Equilibria* 96 (May 1994), pp. 33–50. DOI: 10.1016/0378-3812(94)80086-3. URL: https://doi.org/10.1016/0378-3812(94)80086-3.

[227] Marianne Catté, Claude-Gilles Dussap, and Jean-Bernard Gros. "A physical chemical UNIFAC model for aqueous solutions of sugars". In: *Fluid Phase Equilibria* 105.1 (Mar. 1995), pp. 1–25. DOI: 10.1016/0378-3812(94)02604-y. URL: https://doi.org/10.1016/0378-3812(94)02604-y.

[228] António M. Peres and Eugénia A. Macedo. "Thermodynamic properties of sugars in aqueous solutions: correlation and prediction using a modified UNIQUAC model". In: *Fluid Phase Equilibria* 123.1-2 (Aug. 1996), pp. 71–95. DOI: 10.1016/s0378-3812(96)90013-8. URL: https://doi.org/10.1016/s0378-3812(96)90013-8.

[229] Olga Ferreira, Esteban A Brignole, and Eugenia A Macedo. "Phase equilibria in sugar solutions using the A-UNIFAC model". en. In: *Ind. Eng. Chem. Res.* 42.24 (Nov. 2003), pp. 6212–6222.

[230] Panagiotis Tsavas et al. "Phase equilibrium calculations in aqueous and nonaqueous mixtures of sugars and sugar derivatives with a group-contribution model". en. In: *Ind. Eng. Chem. Res.* 43.26 (Dec. 2004), pp. 8391–8399.

[231] Denise Perozin, Alessandra L. Oliveira, and Fernando A. Cabral. "Modeling of phase equilibria for aqueous solutions of sugars using a cubic equation of state". In: *Journal of*

Food Process Engineering 30.5 (Oct. 2007), pp. 593–606. DOI: 10.1111/j.1745-4530.2007.00132.x. URL: https://doi.org/10.1111/j.1745-4530.2007.00132.x.

[232] Jadwiga Nowak et al. "Adsorption behaviour of sugars versus their activity in single and multicomponent liquid solutions". In: *Journal of Chromatography A* 1216.50 (Dec. 2009), pp. 8697–8704. DOI: 10.1016/j.chroma.2009.01.043. URL: https://doi.org/10.1016/j.chroma.2009.01.043.

[233] R. G. M. van der Sman. "Predicting the solubility of mixtures of sugars and their replacers using the Flory–Huggins theory". In: *Food & Function* 8.1 (2017), pp. 360–371. DOI: 10.1039/c6fo01497f. URL: https://doi.org/10.1039/c6fo01497f.

[234] Lucas T. Paese et al. "Predicting phase equilibrium of aqueous sugar solutions and industrial juices using COSMO-SAC". In: *Journal of Food Engineering* 274 (June 2020), p. 109836. DOI: 10.1016/j.jfoodeng.2019.109836. URL: https://doi.org/10.1016/j.jfoodeng.2019.109836.

[235] Jean Timmermans. *The physico-chemical constants of binary systems in concentrated solutions*. eng. New York: Interscience Publishers, 1959.

[236] Maciej Starzak and Steve Peacock. "Water activity coefficient in aqueous solutions of sucrose - A comprehensive data analysis". In: *Zuckerindustrie. Sugar industry* 122 (Jan. 1997), pp. 380–387.

[237] W. Wagner and A. Pruß. "The IAPWS Formulation 1995 for the Thermodynamic Properties of Ordinary Water Substance for General and Scientific Use". In: *Journal of Physical and Chemical Reference Data* 31.2 (June 2002), pp. 387–535. DOI: 10.1063/1.1461829. URL: https://doi.org/10.1063/1.1461829.

[238] W. F. Giauque and J. W. Stout. "The Entropy of Water and the Third Law of Thermodynamics. The Heat Capacity of Ice from 15 to 273°K." In: *Journal of the American Chemical Society* 58.7 (July 1936), pp. 1144–1150. DOI: 10.1021/ja01298a023. URL: https://doi.org/10.1021/ja01298a023.

[239] G. L. Anderson, Howard Higbie, and Gebhard Stegeman. "The Heat Capacity of Sucrose from 25 to 90°C". In: *Journal of the American Chemical Society* 72.8 (Aug. 1950), pp. 3798–3799. DOI: 10.1021/ja01164a503. URL: https://doi.org/10.1021/ja01164a503.

[240] Robert L. Putnam and Juliana Boerio-Goates. "Heat-capacity measurements and thermodynamic functions of crystalline sucrose at temperatures from 5 K to 342 K." In: *The Journal of Chemical Thermodynamics* 25.5 (May 1993), pp. 607–613. DOI: 10.1006/jcht.1993.1055. URL: https://doi.org/10.1006/jcht.1993.1055.

[241] Gerardo O. Hernández-Segura et al. "Temperature dependence of the heat capacities in the solid state of 18 mono-, di-, and poly-saccharides". In: *The Journal of Chemical Thermodynamics* 41.1 (Jan. 2009), pp. 17–20. DOI: 10.1016/j.jct.2008.08.007. URL: https://doi.org/10.1016/j.jct.2008.08.007.

[242] Edgar Philip Perman and Horace Leonard Saunders. "The vapour pressures of concentrated cane sugar solutions". In: *Transactions of the Faraday Society* 19.July (1923), p. 112. DOI: 10.1039/tf9231900112. URL: https://doi.org/10.1039/tf9231900112.

[243] W. J. Dunning, H. C. Evans, and M. Taylor. "524. The vapour pressures of concentrated aqueous sucrose solutions up to the pressure of 760 mm". In: *Journal of the Chemical Society (Resumed)* (1951), p. 2363. DOI: 10.1039/jr9510002363. URL: https://doi.org/10.1039/jr9510002363.

[244] Walter L. Crider, Richard H. Milburn, and Stephen D. Morton. "THE EVAPORATION AND REHYDRATION OF AQUEOUS SOLUTIONS". In: *Journal of Meteorology* 13.6 (Dec. 1956), pp. 540–547. DOI: 10.1175/1520-0469(1956)013<0540:tearoa>2.0.co;2. URL: https://doi.org/10.1175/1520-0469(1956)013%3C0540:tearoa%3E2.0.co;2.

[245] R. H. Stokes and R. A. Robinson. "Interactions in Aqueous Nonelectrolyte Solutions. I. Solute-Solvent Equilibria". In: *The Journal of Physical Chemistry* 70.7 (July 1966), pp. 2126–2131. DOI: 10.1021/j100879a010. URL: https://doi.org/10.1021/j100879a010.

[246] Nosaibah Ebrahimi and Rahmat Sadeghi. "Osmotic properties of carbohydrate aqueous solutions". In: *Fluid Phase Equilibria* 417 (June 2016), pp. 171–180. DOI: 10.1016/j.fluid.2016.02.030. URL: https://doi.org/10.1016/j.fluid.2016.02.030.

[247] R. A. Robinson and R. H. Stokes. "ACTIVITY COEFFICIENTS IN AQUEOUS SOLUTIONS OF SUCROSE, MANNITOL AND THEIR MIXTURES AT 25°". In: *The Journal of Physical Chemistry* 65.11 (Nov. 1961), pp. 1954–1958. DOI: 10.1021/j100828a010. URL: https://doi.org/10.1021/j100828a010.

[248] Elmira Behboudi, Hemayat Shekaari, and Mohammed Taghi Zafarani-Moattar. "Water Activity in Aqueous Solution of Sucrose in the Presence of Some Deep Eutectic Solvents". In: *Journal of Chemical & Engineering Data* 66.2 (Jan. 2021), pp. 1043–1054. DOI: 10.1021/acs.jced.0c00849. URL: https://doi.org/10.1021/acs.jced.0c00849.

[249] Robert A. Robinson and Donald A. Sinclair. "The Activity Coefficients of the Alkali Chlorides and of Lithium Iodide in Aqueous Solution from Vapor Pressure Measurements". In: *Journal of the American Chemical Society* 56.9 (Sept. 1934), pp. 1830–1835. DOI: 10.1021/ja01324a003. URL: https://doi.org/10.1021/ja01324a003.

[250] Donald G. Archer. "Thermodynamic Properties of the NaCl+H2O System. II. Thermodynamic Properties of NaCl(aq), NaCl2(cr), and Phase Equilibria". In: *Journal of Physical and Chemical Reference Data* 21.4 (July 1992), pp. 793–829. DOI: 10.1063/1.555915. URL: https://doi.org/10.1063/1.555915.

[251] Donald G. Archer. "Thermodynamic Properties of the KCl+H2O System". In: *Journal of Physical and Chemical Reference Data* 28.1 (Jan. 1999), pp. 1–17. DOI: 10.1063/1.556034. URL: https://doi.org/10.1063/1.556034.

[252] Frederick Barry. "A CALORIMETRIC PROCEDURE FOR DETERMINING THE HEATS OF SLOW REACTIONS. II. THE CALORIMETRY OF A SLOW REACTION: THE HEAT OF INVERSION OF SUCROSE BY ACID." In: *Journal of the American Chemical Society* 42.10 (Oct. 1920), [1911]–1945. DOI: 10.1021/ja01455a001. URL: https://doi.org/10.1021/ja01455a001.

[253] B. C. Hendricks et al. "A Modified Vacuum-Walled Adiabatic Calorimeter". In: *The Journal of Physical Chemistry* 34.2 (Feb. 1930), pp. 418–426. DOI: 10.1021/j150308a014. URL: https://doi.org/10.1021/j150308a014.

[254] Danforth P Miller and Juan J de Pablo. "Calorimetric solution properties of simple saccharides and their significance for the stabilization of biological structure and function". en. In: *J. Phys. Chem. B* 104.37 (Sept. 2000), pp. 8876–8883.

[255] Howard Higbie and Gebhard Stegeman. "The Heat of Solution of Sucrose in Water at 25C". In: *Journal of the American Chemical Society* 72.8 (Aug. 1950), pp. 3799–3799. DOI: 10.1021/ja01164a504. URL: https://doi.org/10.1021/ja01164a504.

[256] R.V. Jasra and J.C. Ahluwalia. "Enthalpies and heat capacities of dissolution, apparent molar heat capacities, and apparent molar volumes of some mono-, di-, tri-, and tetrasaccharides in water". In: *The Journal of Chemical Thermodynamics* 16.6 (June 1984), pp. 583–590. DOI: 10.1016/0021-9614(84)90010-7. URL: https://doi.org/10.1016/0021-9614(84)90010-7.

[257] G. Salvetti et al. "Excess energy of polymorphic states or glass over the crystal state by heat of solution measurement". In: *Thermochimica Acta* 285.2 (Aug. 1996), pp. 243–252. DOI: 10.1016/0040-6031(96)02915-2. URL: https://doi.org/10.1016/0040-6031(96)02915-2.

[258] Danchen Gao and J.Howard Rytting. "Use of solution calorimetry to determine the extent of crystallinity of drugs and excipients". In: *International Journal of Pharmaceutics* 151.2 (May 1997), pp. 183–192. DOI: 10.1016/s0378-5173(97)04895-3. URL: https://doi.org/10.1016/s0378-5173(97)04895-3.

[259] Chumakova and Abrosimov. "Sucrose-water Enthalpy of Solution". In: *DETHERM* 36.6 (Aug. 1995), pp. 655–661.

[260] Alfred W. Porter. "The kinetic theory of osmotic pressure". In: *Transactions of the Faraday Society* 13.December (1917), p. 123. DOI: 10.1039/tf9171300123. URL: https://doi.org/10.1039/tf9171300123.

[261] R. B. Vallender and E. P. Perman. "Vapour pressure and heat of dilution.—Part VIII. Heat of dilution of cane sugar in aqueous solution and of urea and calcium chloride in alcoholic solution". In: *Trans. Faraday Soc.* 27.0 (1931), pp. 124–135. DOI: 10.1039/tf9312700124. URL: https://doi.org/10.1039/tf9312700124.

[262] David R Lide. *CRC handbook of chemistry and physics, 84th edition*. en. 84th ed. Boca Raton, FL: CRC Press, June 2003.

[263] R. C. Payn and E. P. Perman. "Vapour pressure and heat of dilution.—Part VI.—Heat of dilution of hydrochloric acid, sodium hydroxide and acetic acid". In: *Trans. Faraday Soc.* 25.0 (1929), pp. 599–610. DOI: 10.1039/tf9292500599. URL: https://doi.org/10.1039/tf9292500599.

[264] Miroslav Leskiv, Carlos E S Bernardes, and Manuel E Minas da Piedade. "A calorimetric system based on the LKB 10700-1 flow microcalorimeter". In: *Measurement Science and Technology* 20.7 (June 2009), p. 075107. DOI: 10.1088/0957-0233/20/7/075107. URL: https://doi.org/10.1088/0957-0233/20/7/075107.

[265] David J. Wales and Jonathan P. K. Doye. "Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms". In: *The Journal of Physical Chemistry A* 101.28 (July 1997), pp. 5111–5116. DOI: 10.1021/jp970984n. URL: https://doi.org/10.1021/jp970984n.

[266] Rainer Storn and Kenneth Price. In: *Journal of Global Optimization* 11.4 (1997), pp. 341–359. DOI: 10.1023/a:1008202821328. URL: https://doi.org/10.1023/a:1008202821328.

[267] J. Kennedy and R. Eberhart. "Particle swarm optimization". In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. IEEE. DOI: 10.1109/icnn.1995.488968. URL: https://doi.org/10.1109/icnn.1995.488968.

[268] Richard Otis. "Uncertainty reduction and quantification in computational thermodynamics". In: *Computational Materials Science* 212 (Sept. 2022), p. 111590. DOI: 10.1016/j.commatsci.2022.111590. URL: https://doi.org/10.1016/j.commatsci.2022.111590.

[269] George L. Nicolaides and Charles A. Eckert. "Optimal Representation of Binary Liquid Mixture Nonidealities". In: *Industrial & Engineering Chemistry Fundamentals* 17.4 (Nov. 1978), pp. 331–340. DOI: 10.1021/i160068a020. URL: https://doi.org/10.1021/i160068a020.

[270] Maximilian Lange. "An artificial neural network model for the unary description of pure substances and its application on the thermodynamic modelling of pure iron". In: *Soft Computing* 24.16 (Jan. 2020), pp. 12227–12239. DOI: 10.1007/s00500-019-04663-3. URL: https://doi.org/10.1007/s00500-019-04663-3.

[271] Lian Yu, Jun Huang, and Karen J. Jones. "Measuring Free-Energy Difference Between Crystal Polymorphs Through Eutectic Melting". In: *J. Phys. Chem. B* 109.42 (2005), pp. 19915–19922. DOI: 10.1021/jp053653g. URL: https://doi.org/10.1021/jp053653g.

[272] Suchart Chongprasert, Shawn A. Knopp, and Steven L. Nail. "Characterization of Frozen Solutions of Glycine". In: *J. Pharm. Sci.* 90.11 (2001), pp. 1720–1728. DOI: 10.1002/jps.1121. URL: https://doi.org/10.1002/jps.1121.

[273] E.Ju. Shalaev et al. "Study of the Phase Diagram Water Fraction of the System Water-Glycine-Sucrose By DTA and X-Ray Diffraction Methods". In: *Thermochim. Acta* 196.1 (1992), pp. 213–220. DOI: 10.1016/0040-6031(92)85021-m. URL: https://doi.org/10.1016/0040-6031(92)85021-m.

Appendix A

Supporting Information for Chapter 3

A.1 Derivation of Differential Solubility Equation (based on Williamson (1944))

Here we present a complete derivation from fundamental thermodyamic arguments for Eq. (3.1) in Chapter 3. We attribute this derivation to [143] and present it here as an additional resource for interested readers.

The Gibbs free energy change for a hypothetical thermodynamic reaction (at equilibrium) is defined as:

$$\Delta G = G^{products} - G^{reactants} \tag{A.1}$$

$$= \Delta H - T\Delta S, \tag{A.2}$$

where ΔH and ΔS are the enthalpy and entropy change associated with the reaction, respectively, and T is temperature. The Gibbs-Helmholtz equation is given as:

$$\frac{d}{dT}\left(\frac{\Delta G}{T}\right) = -\frac{\Delta H}{T^2}, \tag{A.3}$$

which relates the temperature dependence of the Gibbs free energy of the hypothetical process to the enthalpy change of the process. As an aside, we can interpret ΔG further. For the reversible reaction $\mathrm{solute}(s) \leftrightarrow \mathrm{solute}(aq)$, which describes the equilibrium between a pure solid solute and a saturated solution:

$$\Delta G = G^{products} - G^{reactants} \tag{A.4}$$

$$= n_1\mu_1^l - n_1\mu_1^{s,\circ} \tag{A.5}$$

$$\Delta g = \left(\mu_1^{l,\circ} + RT\ln a_1\right) - \mu_1^{s,\circ}, \tag{A.6}$$

where 1 denotes solute, $\mu_1^{s,\circ}$ is the molar Gibbs free energy of the pure solid solute, $\mu_1^{l,\circ}$ is the molar Gibbs free energy of the solute in solution at the chosen reference state, and μ_1^l is the partial molar Gibbs free energy of the solute (i.e. chemical potential) in solution. For a chemical reaction at equilibrium, $\Delta G = 0$, which means $\Delta g = 0$, thus:

$$\Delta g_1^\circ = -RT \ln a_1 \tag{A.7}$$

where a_1 is the solute activity, and we introduce $\Delta g_1^\circ = \mu_1^{l,\circ} - \mu_1^{s,\circ}$ as the standard molar Gibbs free energy for the dissolution process (i.e. the difference between the partial molar Gibbs free energy between the pure solid solute and the solute in solution, in their standard states, which is typically interpretted in terms of the melting properties of the solute). Alternatively, as described by Williamson, the standard Gibbs free energy change for the dissolution equilibrium reaction can be expressed in terms of the equilibrium constant K via:

$$\Delta g_1^\circ = -RT \ln K, \tag{A.8}$$

where $K = a_1$. Following [143] and combining either Eq. (A.7) or Eq. (A.8) with Eq. (A.3) we note that:

$$\frac{d \ln a_1}{dT} = \frac{(h_1^\infty - h_1^{\circ,s})}{RT^2}, \tag{A.9}$$

where $h_1^\infty - h_1^{l,\circ}$ is defined as the standard enthalpy change for the dissolution process (i.e. $\mathrm{solute}(s) \leftrightarrow \mathrm{solute}(aq)$), which is also referred to as the experimentally measurable enthalpy of solution to infinite dilution (noting that we use the infinitely dilute reference state to define a). Further, we note that $a = \gamma x$. Now, we need to define the derivative on the left hand side. Noting that the relationship is only valid along the solubility curve:

$$\left(\frac{d \ln(x_1 \gamma_1)}{dT}\right)_{sat} = \left(\frac{d \ln x_1}{dT}\right)_{sat} + \left(\frac{d \ln \gamma_1}{dT}\right)_{sat}. \tag{A.10}$$

As described by [143], special attention should be given to $\left(\frac{d \ln \gamma_1}{dT}\right)_{sat}$, because it is both a function of temperature and composition, both of which vary as we move along the solubility curve. We can do so by introducing the total derivative:

$$d \ln \gamma_1 = \left(\frac{\partial \ln \gamma_1}{\partial T}\right)_{\ln x} dT + \left(\frac{d \ln \gamma_1}{\partial \ln x_1}\right)_T d \ln x_1, \tag{A.11}$$

which gives (imposing the constraint that differentiation occurs along the solubility curve):

$$\left(\frac{d \ln \gamma_1}{dT}\right)_{sat} = \left(\frac{\partial \ln \gamma_1}{\partial T}\right)_{\ln x} + \left(\frac{d \ln \gamma_1}{\partial \ln x}\right)_T \left(\frac{d \ln x_1}{dT}\right)_{sat}. \tag{A.12}$$

Here, we note that:

$$\left(\frac{\partial \ln \gamma_1}{\partial T}\right)_{\ln x} = -\frac{(\bar{h}_1 - h_1^\infty)}{RT^2}, \tag{A.13}$$

which is often termed the 'relative partial molar heat content of the solute'. Coupled with the standard enthalpy change for the dissolution process:

$$\Delta h_1^{soln,s} = \frac{(\bar{h}_1 - h_1^{\infty})}{RT^2} + \frac{(h_1^{\infty} - h_1^{s,\circ})}{RT^2} \tag{A.14}$$

$$= \frac{(\bar{h}_1 - h_1^{s})}{RT^2}. \tag{A.15}$$

Combining all of the above, we get:

$$\frac{d\ln x_1}{dT} + \left(\frac{\partial \ln \gamma_1}{\partial \ln x_1}\right)_T \left(\frac{d\ln x_1}{dT}\right)_{sat} = \frac{(\bar{h}_1 - h_1^{s})}{RT^2} \tag{A.16}$$

$$\frac{d\ln x_1}{dT} = \frac{1}{RT^2}\left[\frac{(\bar{h}_1 - h_1^{s})}{1 + \partial \ln \gamma_1 / \partial \ln x_1}\right]. \tag{A.17}$$

A.2 Solubility Fits for Estimating Solubility Ratio

Certain solubility data reports contain measurements for both α and γ-glycine. In theory, these can be processed to determine the solubility ratio of polymorphs at various temperatures. In some cases, measurements are for both polymorphs are reported at the same temperature; however, in others, they are reported at different temperatures.

For data sets reporting data at different temperatures for each polymorph, the following solubility model was used to interpolate a uniform set of solubility, thus giving estimates for the solubility ratio:

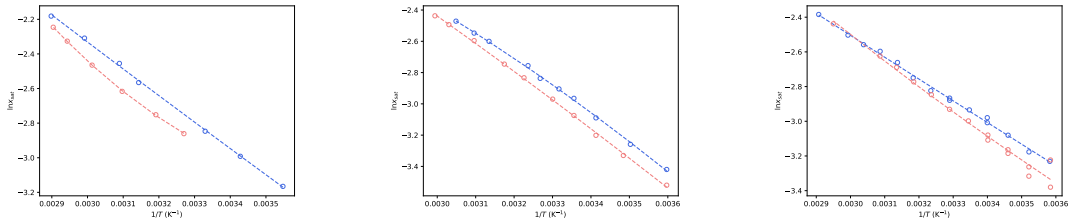
$$\ln x = \frac{-a}{R}\frac{1}{T} + \frac{b}{R}\ln T + c \tag{A.18}$$

Table A.1: α and γ-glycine parameters for Eq. (A.18) regression

	Park et. al		Yang et. al		Igarashi et. al	
Parameter	α	γ	α	γ	α	γ
$a(\times 10^3)$	1.213	6.401	4.435	2.858	1.141	0.144
b	1.041	24.908	−8.911	−3.773	0.371	4.220
c	−4.738	−166.390	62.676	25.756	−1.236	−26.583

A.3 Bouchard Solubility Ratio Data and Interpretation

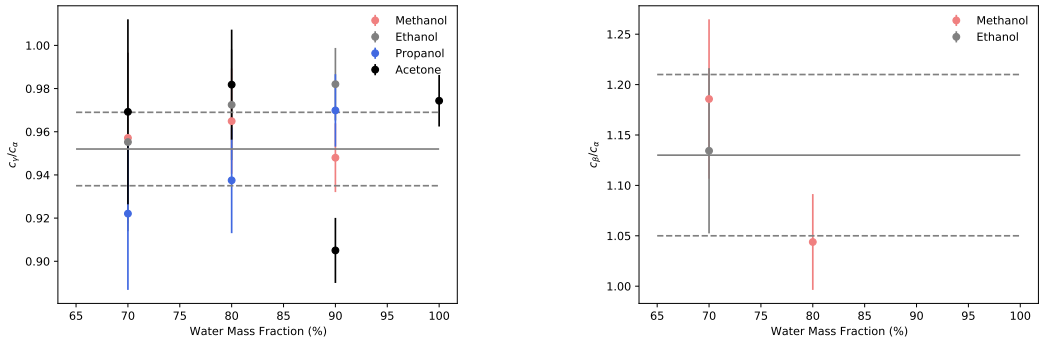
Bouchard et al., measured the solubility of α, β and γ-glycine in aqueous solutions of varying composition containing methanol, ethanol, 2-propanol or acetone at 310 K. β-glycine solubility in



(a) Park et. al. (b) Yang et. al. (c) Igarashi et. al.

Figure A.1: Fits to selected literature solubility data (blue; α, red; γ, open circles; data, dashed lines; fits)

pure water wasn't measured due to assumed rapid re-crystallisation to α. As shown in Fig. A.2, extrapolating the solubility ratio (c_γ/c_α) measured in various mixed solvent systems (above 70 wt% water) gives a reasonable estimate for the solubility ratio in pure water. Thus, the solubility ratio (c_β/c_α in pure water was estimated using data from mixed solvent solubility measurements—shown graphically in Fig. A.2.



(a) γ-α (b) β-α

Figure A.2: Selected Bouchard solubility ratios with estimated uncertainties (points; solubility ratios (note: c measured in g glycine / g solution, grey solid line; expected solubility ratio, grey dashed lines; $\pm\sigma^*$)

The uncertainty on ratio estimates derived from Bouchard anti-solvent solubility data was estimated using Uncertainties python package. For γ-α, the solubility ratio uncertainty estimate was calculated using uncertainty data provided in the article (i.e. $\pm$ 0.002 g/gsoln) and translated into mole fraction ratios. For the purpose of estimation, we apply the same uncertainty to the 'average' solubility ratio for β-α.

A.4 Miscellaneous Tabulated Literature Data

In this section, literature data used in various figures throughout the associated manuscript are tabulated for a complete record. This includes: α-γ cross-over temperatures and α, β and γ-glycine eutectic temperatures.

Table A.2: Literaure α-γ cross-over temperatures

Reference	Temperature (K)
[137]	452.14
[105]	441.55–473.15
[138]	438.15 ± 5
[12]	450.15
[139] (via [142])	438
[140]	462.15
[141]	441.15 & 445.15
[142]	425–450

Table A.3: Literaure eutectic temperature and composition estimates (adapted from a review by [130])

Reference	α	β	γ
[130]	-2.8 ± 0.1	-3.6 ± 0.1	-2.8 ± 0.1
[271]	-3.6	-3.8	-2.9
[272]		-3.6	-2.7
[273]		-3.6 ± 0.1	

A.5 Freezing Point Depression

In general, the presence of solute in solvent leads to 'freezing point depression' (i.e. a decrease in the temperature at which a solvent freezes). It should be noted that freezing point depression occurs while the concentration of solute in solution is below the eutectic composition for the solute/solvent system. The freezing point of glycine-water solutions (or related properties from which the freezing point can be derived) has been measured by various authors [131, 132, 133,

134, 135] and presented in Fig. A.3. It should be noted that reports of water osmotic coefficients at the freezing point were converted to freezing point temperatures using:

$$\Delta T \approx \frac{\Phi}{1.858} m_g,$$

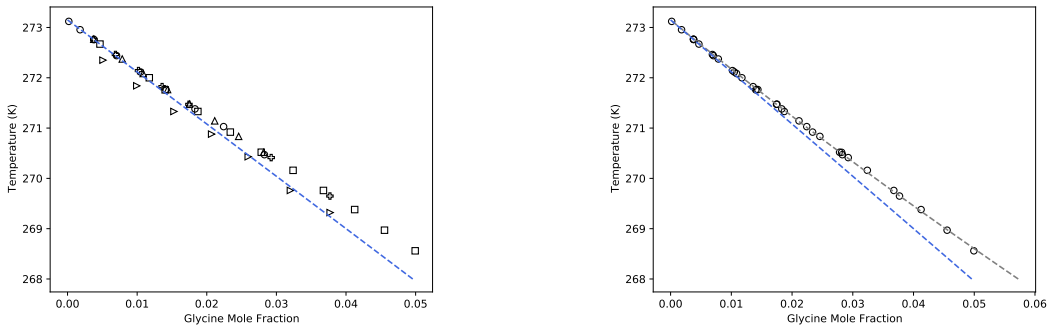
where $\Delta T = 0°C - T_{fp}$. In addition, the freezing point depression was predicted using:

$$\ln x_w = \frac{\Delta H^{fus}}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right),$$

where $\Delta H^{fus} = 6010$ J/mol and $T_m = 273.15$ K (i.e. the melting properties of pure ice). From Fig. A.3, four of the data sets are consistent, while the data reported by Shimoyamada et. al., appears to be inconsistent. The consistency of the other four data sets is further supported by modeling work of Rowland [80]. The following model was used to correlate the consistent data:

$$\ln x_w = -A\left(\frac{1}{T} - \frac{1}{T_m}\right) + B \ln\left(\frac{T}{T_m}\right),$$

where $A = 10115.62$ K and $B = -34.29$, and is shown in Fig. A.3.



(a) All data

(b) Selected data with best fit

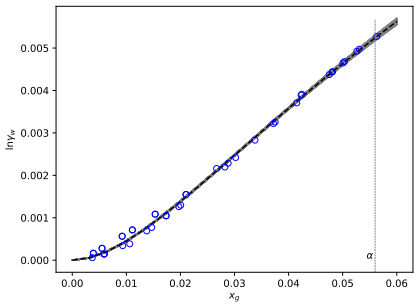
Figure A.3: Freezing point depression data (symbols; experimental data, blue dahsed line; prediction from ice melting properties, grey dashed line; best fit)

A.6 Sensitivity Analysis

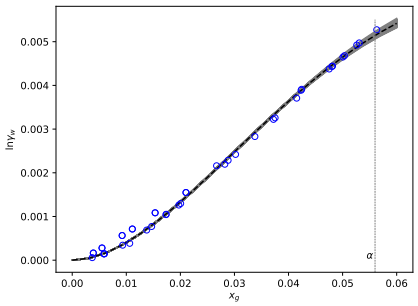
In the associated manuscript, we present a range of predictions for the aqueous solubility of α-glycine, based on a sensitivity analysis of the underlying thermodynamic data used to regress excess Gibbs free energy models, as well as sensitivity to the form of the chosen models. Here we illustrate the aforementioned sensitivity, for all levels of assumed measurement uncertainty for

$\ln \gamma_w$, $\Delta H^{dil}/n_g$ and c_p^{soln}. Figures are grouped together based on level of assumed measurement uncertainty (i.e. σ^*, $2\sigma^*$ and $5\sigma^*$)—as described in the associated manuscript. Finally, we present α-glycine aqueous solubility predictions based on underlying measurement uncertainty to each themrodynamic quantity (i.e. $\ln \gamma_w$, ΔH^{dil} and c_p^{soln}) at each uncertainty level, for the Scatchard-Hildebrand model.

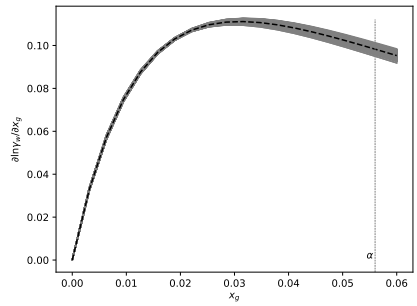
A.6.1 Uncertainty Level: σ^*



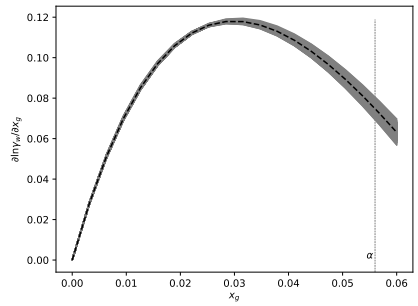
(a) Scatchard-Hildebrand



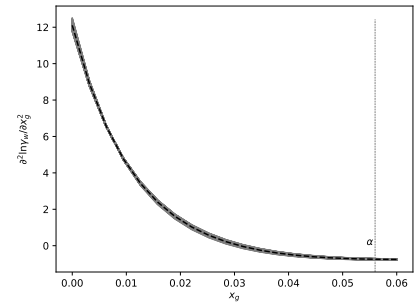
(b) Scatchard-Hildebrand-Flory-Huggins



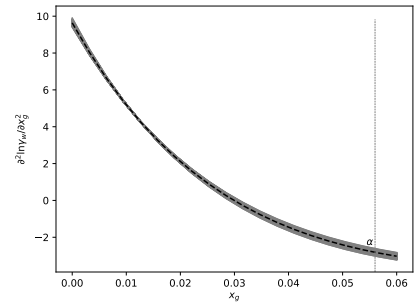
(c) Scatchard-Hildebrand



(d) Scatchard-Hildebrand-Flory-Huggins

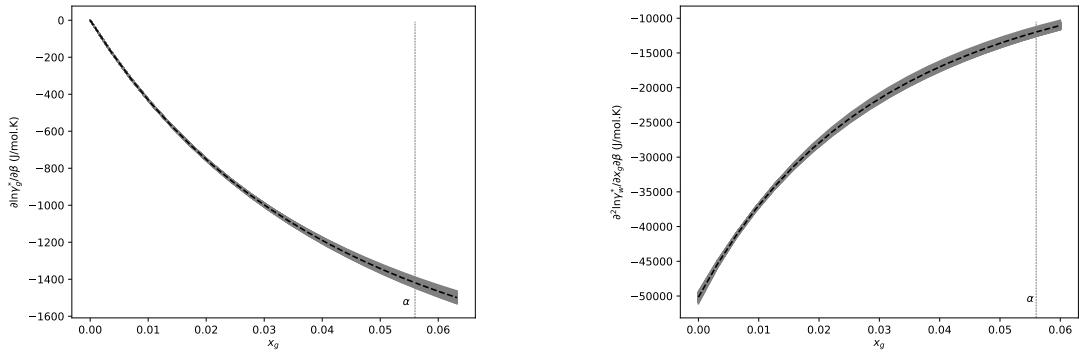


(e) Scatchard-Hildebrand



(f) Scatchard-Hildebrand-Flory-Huggins

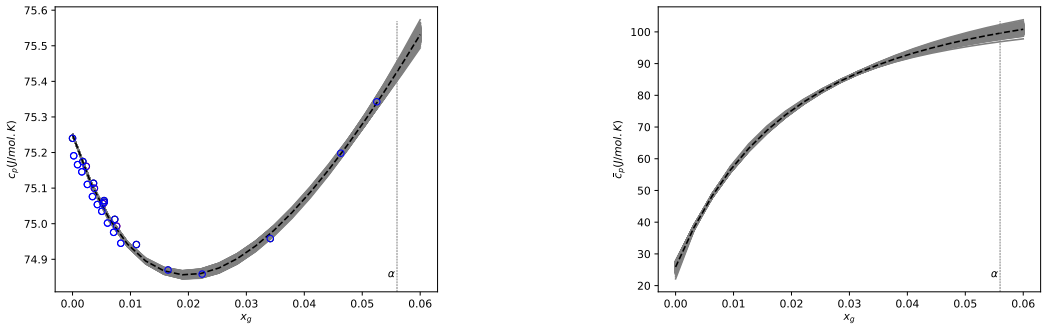
Figure A.4: Sensitivity of water activity and associated partial derivatives in glycine-water mixtures at 298.15 K to assumed measurement uncertainty of σ^*. (a)–(b) show $\ln\gamma_w$ (where blue open circles are selected data from literature as discussed in the associated manuscript), (b)–(c) show $\partial\ln\gamma_w/\partial x_g$ and (d)–(e) show $\partial^2\ln\gamma_w/\partial x_g^2$. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K



(a) $\partial \ln \gamma_w^* / \partial \beta$

(b) $\partial^2 \ln \gamma_w^* / \partial x_g \partial \beta$

Figure A.5: Sensitivity of $\partial \ln \gamma_g^* / \partial \beta$ and $\partial^2 \ln \gamma_g / \partial x_g \partial \beta$ at 298.15 K to assumed measurement uncertainty in $\Delta H^{dil} / n_g$. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K

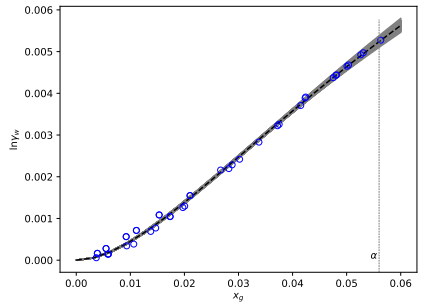


(a) c_p

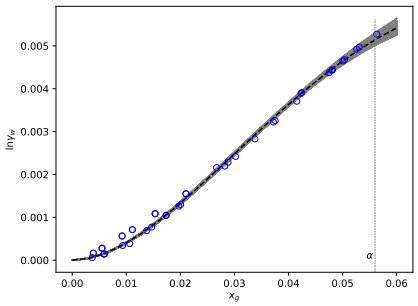
(b) $\bar{c}_{p,g}$

Figure A.6: Sensitivity of glycine-water solution heat capacity and glycine partial molar heat capacity in glycine-water mixtures at 298.15 K to assumed measurement uncertainty in c_p^{soln}. Blue open circles indicate selected literature data. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K

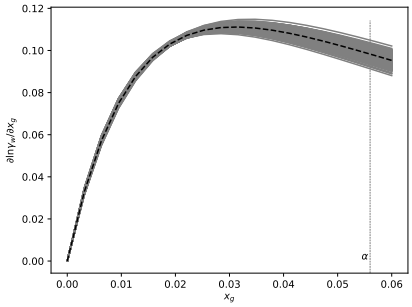
A.6.2 Uncertainty Level: 2 σ^*



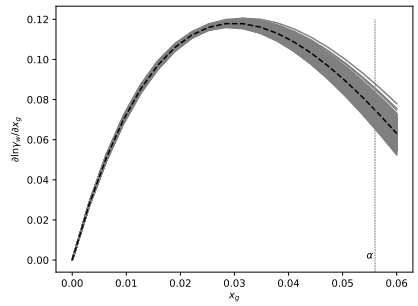
(a) Scatchard-Hildebrand



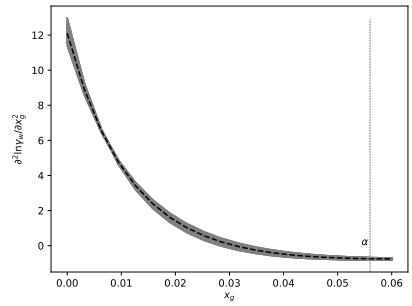
(b) Scatchard-Hildebrand-Flory-Huggins



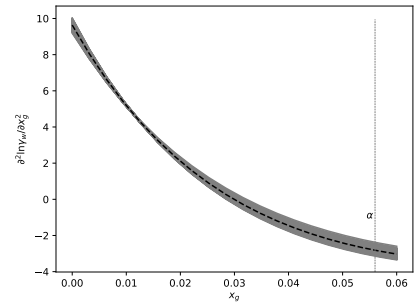
(c) Scatchard-Hildebrand



(d) Scatchard-Hildebrand-Flory-Huggins

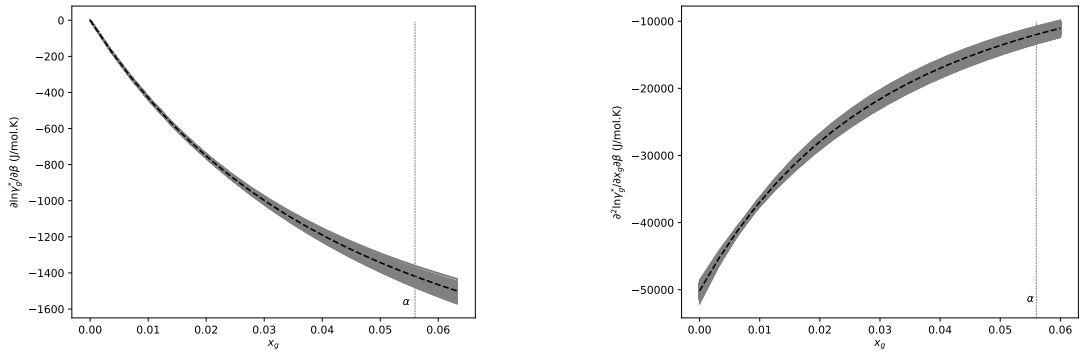


(e) Scatchard-Hildebrand



(f) Scatchard-Hildebrand-Flory-Huggins

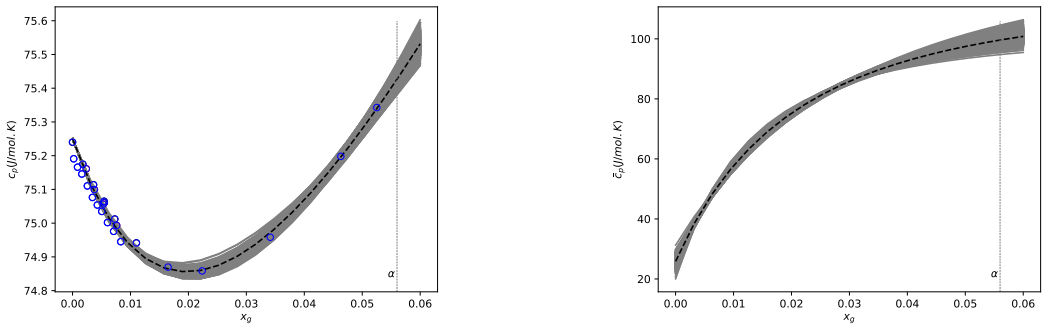
Figure A.7: Sensitivity of water activity and associated partial derivatives in glycine-water mixtures at 298.15 K to assumed measurement uncertainty of $2\sigma^*$. (a)–(b) show $\ln \gamma_w$ (where blue open circles are selected data from literature as discussed in the associated manuscript), (b)–(c) show $\partial \ln \gamma_w / \partial x_g$ and (d)–(e) show $\partial^2 \ln \gamma_w / \partial x_g^2$. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K



(a) $\partial \ln \gamma_w^* / \partial \beta$

(b) $\partial^2 \ln \gamma_w^* / \partial x_g \partial \beta$

Figure A.8: Sensitivity of $\partial \ln \gamma_g^* / \partial \beta$ and $\partial^2 \ln \gamma_g / \partial x_g \partial \beta$ at 298.15 K to assumed measurement uncertainty in $\Delta H^{dil} / n_g$. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K

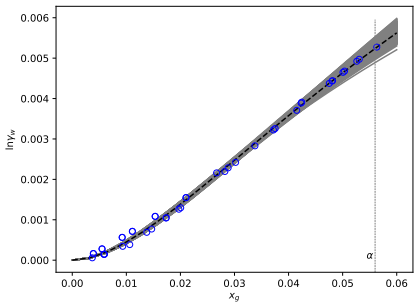


(a) c_p

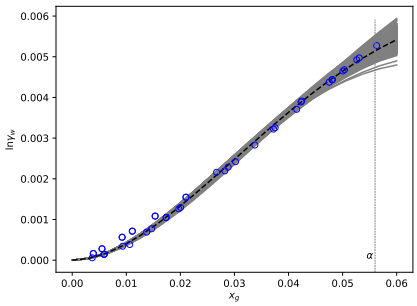
(b) $\bar{c}_{p,g}$

Figure A.9: Sensitivity of glycine-water solution heat capacity and glycine partial molar heat capacity in glycine-water mixtures at 298.15 K to assumed measurement uncertainty in c_p^{soln}. Blue open circles indicate selected literature data. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K

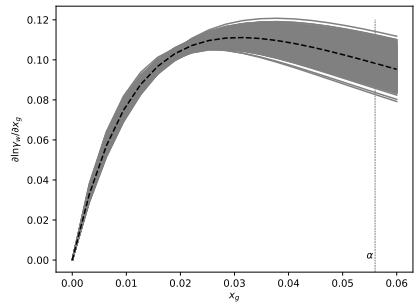
A.6.3 Uncertainty Level: 5 σ^*



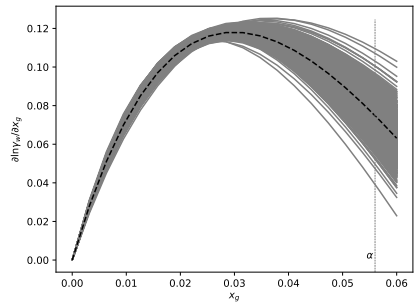
(a) Scatchard-Hildebrand



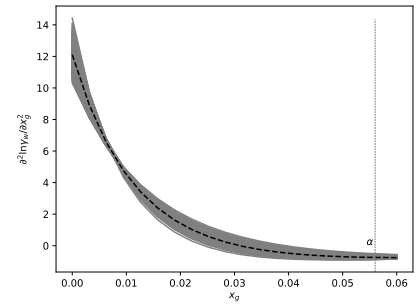
(b) Scatchard-Hildebrand-Flory-Huggins



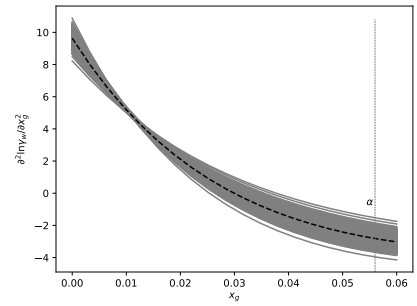
(c) Scatchard-Hildebrand



(d) Scatchard-Hildebrand-Flory-Huggins



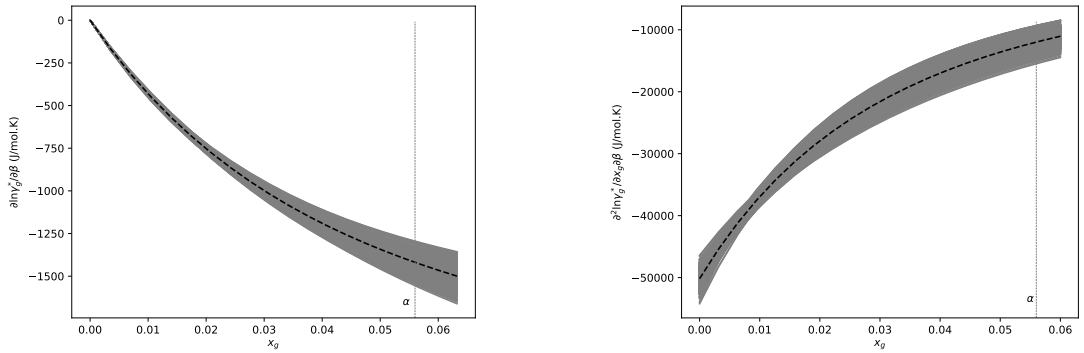
(e) Scatchard-Hildebrand



(f) Scatchard-Hildebrand-Flory-Huggins

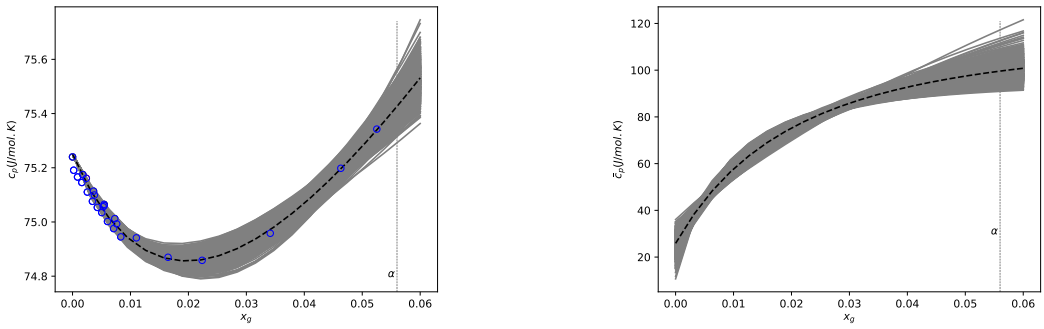
Figure A.10: Sensitivity of water activity and associated partial derivatives in glycine-water mixtures at 298.15 K to assumed measurement uncertainty of $5\sigma^*$. (a)–(b) show $\ln\gamma_w$ (where blue open circles are selected data from literature as discussed in the associated manuscript), (b)–(c) show $\partial\ln\gamma_w/\partial x_g$ and (d)–(e) show $\partial^2\ln\gamma_w/\partial x_g^2$. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K

A.6.4 α-Glycine Solubility Predictions



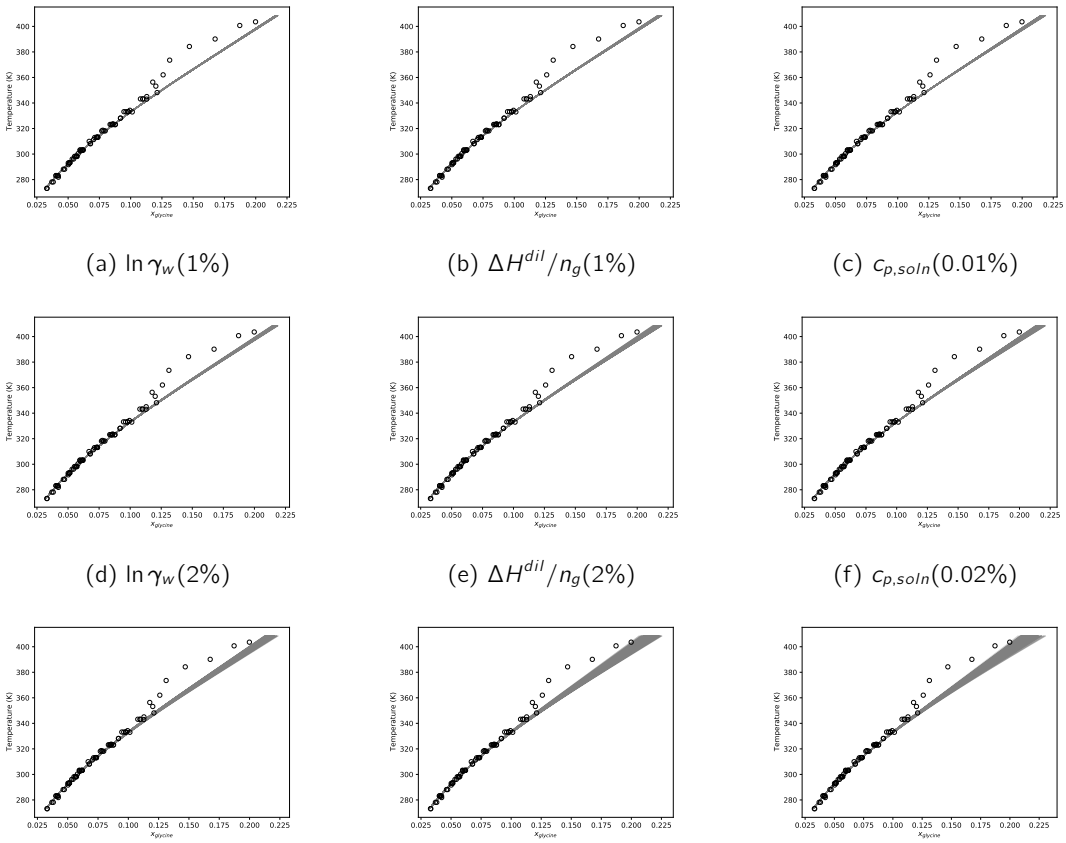
(a) $\partial \ln \gamma_w^* / \partial \beta$ (b) $\partial^2 \ln \gamma_w^* / \partial x_g \partial \beta$

Figure A.11: Sensitivity of $\partial \ln \gamma_g^* / \partial \beta$ and $\partial^2 \ln \gamma_g / \partial x_g \partial \beta$ at 298.15 K to assumed measurement uncertainty in $\Delta H^{dil} / n_g$. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K



(a) c_p (b) $\bar{c}_{p,g}$

Figure A.12: Sensitivity of glycine-water solution heat capacity and glycine partial molar heat capacity in glycine-water mixtures at 298.15 K to assumed measurement uncertainty in c_p^{soln}. Blue open circles indicate selected literature data. Vertical dashed line indicates α-glycine aqueous solubility at 298.15 K



(a) $\ln \gamma_w$(1%) (b) $\Delta H^{dil}/n_g$(1%) (c) $c_{p,soln}$(0.01%)

(d) $\ln \gamma_w$(2%) (e) $\Delta H^{dil}/n_g$(2%) (f) $c_{p,soln}$(0.02%)

(g) $\ln \gamma_w$(5%) (h) $\Delta H^{dil}/n_g$(5%) (i) $c_{p,soln}$(0.05%)

Figure A.13: α-glycine solubility prediction sensitivity to independent sources of thermodynamic data using the Scatchard-Hildebrand model (black circles; direct solubility measurements, grey fan; solubility predictions)

A.7 Tabulated Solubility Predictions

In the associated manuscript, we graphically present various predictions for the aqueous solubility of α, β and γ-glycine across a broad range of temperatures. To make re-use of this data simpler, here we present tabulated values corresponding to each prediction at 5 K increments between 263.15 and 398.15 K, where the solubility is given in terms of glycine mole fraction.

Table A.4: Solubility predictions for α, β and γ-glycine in water at 5 K increments between 263.15 and 398.15 K derived from predictions described in the associated manuscript. Solubility reported in glycine mole fraction. SH and SHFH denote predictions based on Scatchard-Hildebrand and Scatchard-Hildebrand-Flory-Huggins excess Gibbs free energy models, respectively.

	α		β		γ	
T (K)	SH	SHFH	SH	SHFH	SH	SHFH
263.15	0.025	0.024	0.034	0.034	0.023	0.022
268.15	0.028	0.028	0.039	0.039	0.026	0.026
273.15	0.032	0.032	0.044	0.044	0.029	0.029
278.15	0.036	0.036	0.049	0.049	0.033	0.033
283.15	0.041	0.041	0.055	0.055	0.038	0.038
288.15	0.046	0.046	0.061	0.061	0.042	0.042
293.15	0.051	0.051	0.067	0.067	0.047	0.047
298.15	0.056	0.056	0.074	0.073	0.052	0.052
303.15	0.062	0.061	0.081	0.079	0.057	0.057
308.15	0.067	0.067	0.088	0.086	0.063	0.063
313.15	0.074	0.073	0.096	0.093	0.069	0.068
318.15	0.080	0.079	0.104	0.099	0.075	0.074
323.15	0.086	0.085	0.112	0.106	0.081	0.080
328.15	0.093	0.091	0.120	0.113	0.088	0.086
333.15	0.100	0.097	0.129	0.120	0.095	0.092
338.15	0.107	0.103	0.138	0.127	0.102	0.098
343.15	0.115	0.109	0.146	0.134	0.109	0.105
348.15	0.122	0.116	0.156	0.141	0.116	0.111
353.15	0.129	0.122	0.165	0.148	0.123	0.117
358.15	0.137	0.128	0.174	0.155	0.131	0.124
363.15	0.145	0.134	0.184	0.161	0.139	0.130
368.15	0.152	0.141	0.193	0.168	0.146	0.136
373.15	0.160	0.147	0.203	0.175	0.154	0.143
378.15	0.168	0.153	0.213	0.181	0.162	0.149
383.15	0.176	0.159	0.223	0.188	0.170	0.155
388.15	0.184	0.165	0.232	0.195	0.178	0.162
393.15	0.192	0.171	0.242	0.201	0.186	0.168
398.15	0.200	0.177	0.252	0.207	0.194	0.174

Appendix B

Mullin Aqueous Solubility Analysis

Table B.1: Compounds used in Chapter 4 solubility analysis broken down by chemical structure.

Type	Compound	Curvature
Amino Acid	Alanine (D)	+
	Alanine (DL)	+
	Glutamic Acid	+
	Glycine	−
Sugar	Glucose	−
	Maltose	+
	Lactose	+
	Mannitol (D)	+
	Raffinose	−
	Sucrose	+
Aromatic	o-Aminophenol	−
	p-Aminophenol	−
	Benzoic Acid	+
	o-Hydroxybenzoic Acid	+
	m-Hydroxybenzoic Acid	+
	p-Hydroxybenzoic Acid	+
	o-Phthalic Acid	+
	Picric Acid	+
	Salicylic Acid	+
	Hydroquinone	+
	Resorcinol	−
Acid	Adipic Acid	+
	Citric Acid	−
	Fumaric Acid	+
	Maleic Acid	−
	Malic Acid (DL)	+
	Oxalic Acid	−
	Succinic Acid	+
	Tartaric Acid	+
	Tartaric Acid (racemic)	−

Other	Dicyandiamide	+
	Melamine	+
	Pentaerythritol	+
	Succinimide	−
	Taurine	−
	Thiourea	−
	Urea	−
	Uric Acid	−

Table B.2: Parameter and RSS values derived from fitting aquous solubility data reported in [6]

Compound	A	B	C	RSS
Adipic Acid	-475.04	17034.54	72.31	0.00056
Alanine (D)	-37.34	717.95	5.53	3.76e-07
Alanine (DL)	-44.10	863.65	6.63	5.62e-07
o-aminophenol	6.28	-1126.29	-1.43	2.34e-05
p-aminophenol	-302.78	11741.44	45.19	0.00069
benzoic acid	-296.89	10454.89	44.62	9.80e-05
citric acid	31.87	-2744.97	-4.34	0.00052
dicyandiamide	-28.85	-2346.21	5.61	0.00063
fumaric acid	-172.85	4483.81	26.47	5.41e-05
glucose (dextrose)	35.73	-3679.51	-4.51	0.0015
glutamic acid (D)	-145.63	3375.90	22.37	4.39e-05
glycine	55.96	-4035.72	-7.95	4.98e-05
hydroquinone	-352.35	12990.92	53.44	0.00097
o-hydroxybenzoic acid	-377.32	14025.49	56.55	7.89e-05
m-hydroxybenzoic acid	-227.16	6583.35	34.83	0.00030
p-hydroxybenzoic acid	-468.84	17148.90	70.95	0.00026
lactose	-146.04	4375.03	22.27	2.52e-05
maleic acid	26.58	-3035.08	-3.27	0.00042
malic acid (DL)	-97.01	3090.52	14.88	0.00025
malonic acid	-43.44	891.53	6.83	0.00061
maltose	-182.40	6677.26	27.53	0.00010
mannitol (D)	-91.37	1644.36	14.39	0.00010
melamine	-30.02	-2262.22	5.27	0.00033
oxalic acid	66.48	-6308.56	-8.62	2.21e-05
pentaerythritol	-88.54	1243.94	13.98	9.44e-05
o-phthalic acid	-157.65	3571.59	24.30	5.84e-05
picric acid	-195.10	7035.11	28.91	0.00028
raffinose	402.54	-22636.64	-58.19	0.0020
resorcinol	88.48	-5876.50	-12.36	0.0010
salicylic acid	-377.32	14025.49	56.55	7.89e-05
succinic acid	-11.57	-2895.15	2.97	0.00015
succinimide	346.08	-18979.29	-50.04	0.0057

sucrose	-75.77	2752.43	11.27	2.09e-05
tartaric acid (D or L)	-33.22	703.51	5.08	8.77e-05
tartaric acid (racemic)	102.76	-7361.33	-14.34	0.00033
taurine	153.18	-9534.90	-22.01	2.41e-05
thiourea	74.73	-6761.61	-9.72	0.00061
urea	14.33	-1921.17	-1.62	0.00026
uric acid	125.36	-9230.07	-18.64	0.00014

Appendix C

Initial Paramater Values

Table C.1: Redlich-Kister initial parameter values and for glycine-water fits presented in Chapter 4

Model	A_0	B_0	C_0	D_0	
1	0				
2	0	0			
3	0	0	0		
4	0	0	0	0	
Model	A_1	B_1	C_1	D_1	
1	0				
2	0	0			
3	0	0	0	0	
4	1000	1000	1000	-1000	
Model	A_2	B_2	C_2	D_2	$c_{p,1}^{s,o}$
1	0				700
2	-600	-500			700
3	500	-500	500		700
4	500	-500	500	-500	700

Table C.2: Redlich-Kister initial parameter values and for sucrose-water fits presented in Chapter 4

Model	A_0	B_0	C_0	D_0	
1	0				
2	0	0			
3	0	0	0		
4	0	0	0	0	
Model	A_1	B_1	C_1	D_1	
1	0				
2	0	0			
3	10	-10	10	0	
4	10	-10	10	-10	
Model	A_2	B_2	C_2	D_2	$c_{p,1}^{s,\circ}$
1	0				700
2	-600	-500			700
3	500	-500	500		700
4	500	-500	500	-500	700

Appendix D

Water Vapor Pressure Analysis for Sucrose-water Liquid Mixtures

In this appendix, vapor pressure measurements for water vapor above sucrose-water liquid mixtures are modelled to estimate the differential as discussed in Chapter 5.

D.1 Vapor Pressure Data

From the data review described in Chapter 5, independent reports of vapor pressure measurements at including data from 313.15 K to 373.15 K were identified [168, 242, 164, 150] and the resulting vapor pressures are shown in Fig. D.1.

From the vapor pressure data, the activity coefficient in water was determined using modified Raoult's Law (see Eq. (5.7)). Crucially, this requires an estimate for the pure water saturation vapor pressure, which is assumed to be known through standard reference models based on the Antoine equation. Using parameters for the Antoine equation (see Chapter 5), the water activity coefficient was estimated and illustrated in Fig. D.2.

The water activity coefficient was also estimated using pure water saturated vapor pressure estimated by each of the experimental methodologies, to capture any systematic bias. Data reported in [168] and [164] includes measurements for pure water at the associated temperature. For [242] and [150], the required values were estimated by extrapolation to zero sucrose mole fraction using a second-order polynomial implemented in Excel. Using dataset-specific estimates for the pure water vapor pressure, the water activity coefficient was estimated and illustrated in Fig. D.3.

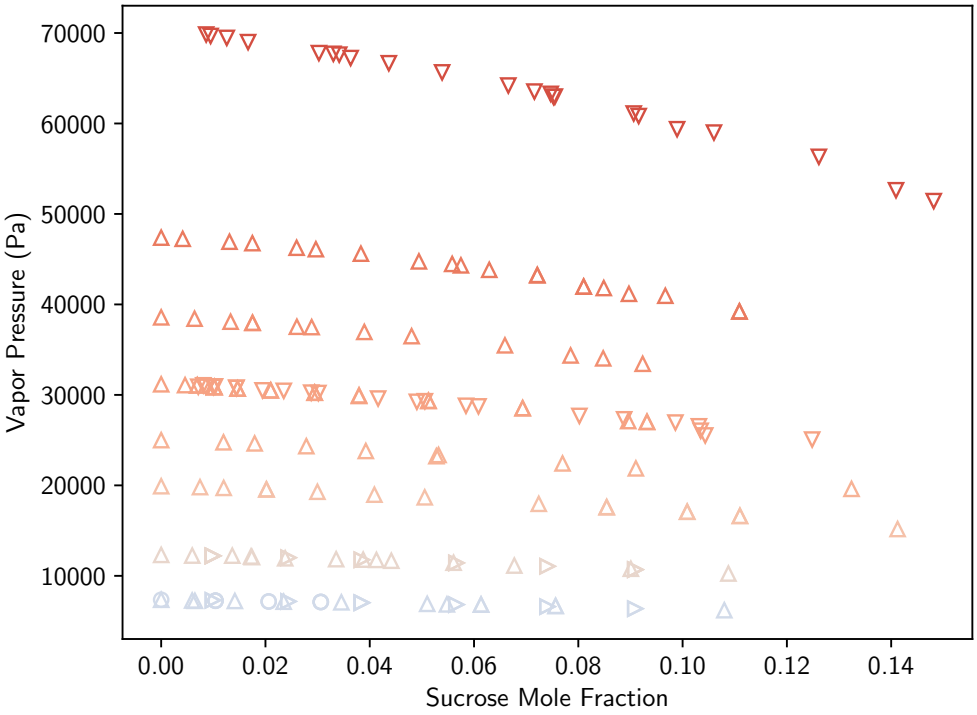


Figure D.1: Vapor pressure measurements reported in literature between 313.15 and 373.15 K as a function of sucrose composition.

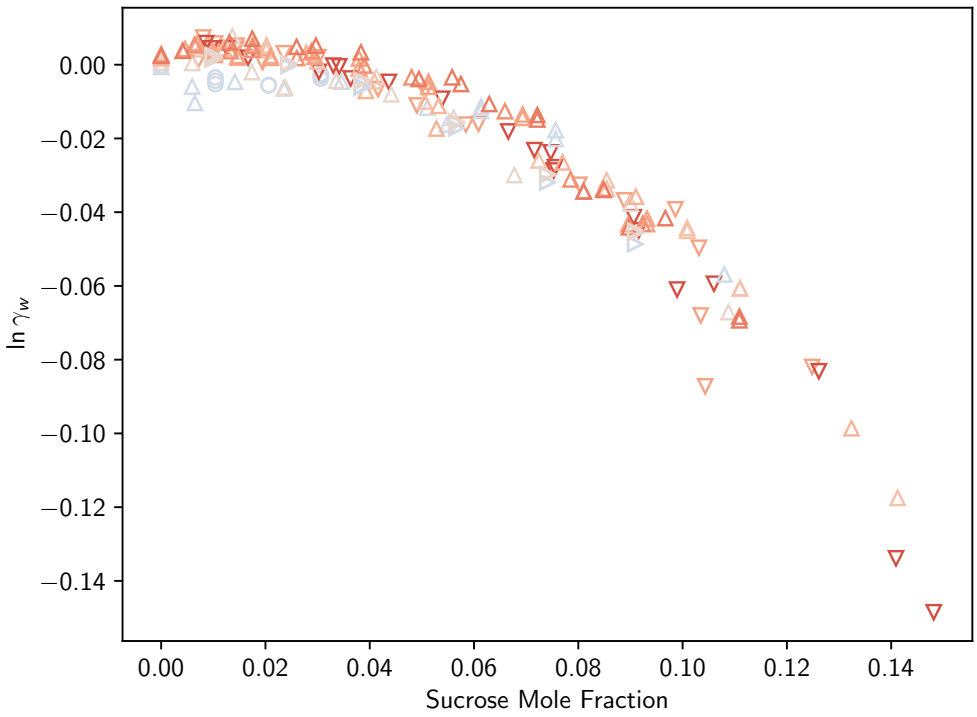


Figure D.2: Water activity coefficient (defined relative to Raoult's Law) between 313.15 and 373.15 K as a function of sucrose composition.

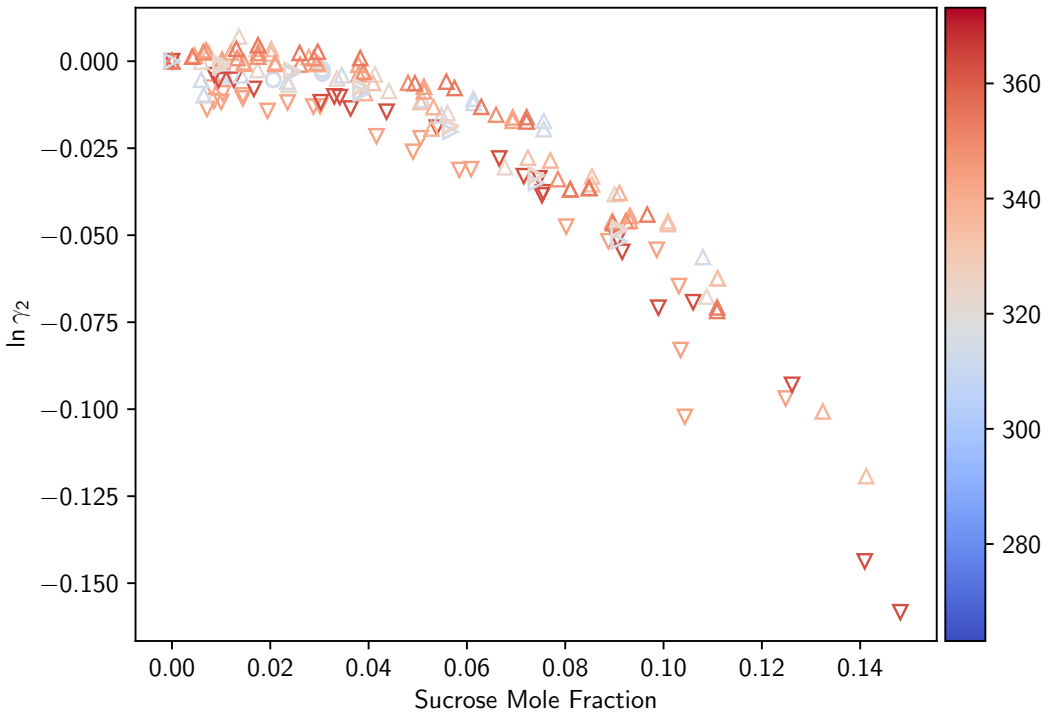


Figure D.3: Water activity coefficient (defined relative to Raoult's Law) between 313.15 and 373.15 K as a function of sucrose composition.